UNIVERSITY OF CALIFORNIA

Los Angeles

Sampling Strategies for Transport Parameter Identification

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Civil Engineering

by

Theodore G. Cleveland

1989

The dissertation of Theodore G. Cleveland is approved.

John A. Dracup

John A. Dracup

Douglas A. Mackay

Douglas A. Mackay

J B Neethling

Johannes B. Neethling

Donald Ylvisaker

Donald Ylvisaker

William W-G. Yeh

William W-G. Yeh, Committee Chair

University of California, Los Angeles

1989

This thesis is dedicated to

Fallen Comrades

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Professor William W-G. Yeh for his support and advice. I would like to thank the rest of my committee, Professors John Dracup, Douglas Mackay, J.B. Neethling and Donald Ylvisaker for their advice and encouragement.

I would like to thank my office mates, Klaus, Ted, Karen, Tracy, James, Mario, Roger, Ben Jei, Jen Min, Don and Uzo for the many hours of conversation and coffee.

Lastly, I would like to thank my mom, dad and sister for putting up with me all along.

The contents of this dissertation were developed under a grant from the Department of the Interior, U.S. Geological Survey (Award No. 14-08-0001-G1499). However those contents do not necessarily represent the policy of the agency, and you should not assume endorsement by the Federal Government.

VITA

November 21, 1959	Born, Larkspur, California
1976-1980	Clinical Medical Specialist, U.S. Army
1986	B.S. Environmental Resources Engineering, Humboldt State University Arcata, California
1986-1987	Chancellors Fellow, Civil Engineering Dept. University of California Los Angeles, California
1987	M.S. Civil Engineering Department University of California Los Angeles, California
1987-1988	Teaching Assistant, Civil Engineering Department University of California Los Angeles, California
1988-1989	Research Assistant, Civil Engineering Department University of California Los Angeles, California

Presentations

Cleveland, T.G. and Yeh, W.W-G., *Sampling Strategies for Transport Parameter Estimation*, American Geophysical Union Fall Meeting, San Francisco, California, December 6-11, 1988.

ABSTRACT OF THE DISSERTATION

Sampling Strategies for Transport

Parameter Identification

by

Theodore G. Cleveland

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 1989

Professor William W-G. Yeh, Committee Chair

This research investigates sampling strategies for a groundwater transport problem, that yield data, which when used in an estimation scheme gives reliable estimates. Reliability is measured using the weighted trace of the information matrix, which is constructed from model sensitivities. The problem is formulated as a constrained binary integer program which is solved using a polynomial time approximation algorithm.

One and two dimensional examples are presented which show that the measure used gives reasonable designs. A method is developed to investigate the range of parameters for which a design remains optimal. Results indicate that design for joint estimation of parameters is multiobjective in nature, designs remain optimal for only a narrow range of parameter settings, and the marginal decrease in information is dramatic when a suboptimal design is used.

1. INTRODUCTION

Management, remediation, and preservation of groundwater systems is a major environmental concern facing society. Water that is collected, stored and transported in the ground provides a major source of water supply. Management of this resource relies on the ability to estimate the effects of stresses placed on the system. Often numerical simulation models are used to predict the movement of groundwater and pollutants carried by the water.

Aquifer response prediction is quite complex. The physical structure of the aquifer system is unobservable in nature. Since the system cannot be observed directly, the system characteristics must be inferred by observing the concentrations and hydraulic head at a few observation wells. The observations are then used to estimate the parameters imbedded in the governing equations that describe the behavior of the system. Collecting these observations can be expensive therefore it is desirable to take observations at points in the aquifer that will yield the most information for a given cost.

In constructing a sampling network, decisions as to the placement of injection and observation sites, associated injection rates, and observation frequencies must be made. One wants to make these decisions to collect a sufficient quantity and quality of data in order to minimize the prediction error. This is an experimental design problem. A method that chooses a design that outperforms all other candidate designs is an optimal experimental design method.

In the design of experiments two approaches are used. The first is one shot experimental design, where the design will be specified once without updating. The second is a sequential design where the results from one experiment are used to construct a design for the next. Each method has some attractiveness, but the choice of approach depends on the consequences of failure. If the consequences of failure are costly, then a one shot robust design is probably preferred. If the consequences of failure are lower, then the sequential approach, which is ultimately cheaper, is probably preferred. In either case the concepts of optimal design are the same. One can even consider sequential design as a series of one shot designs.

The concept of optimal designs is relatively simple. What is required is a way of computing designs without actually running an experiment. In groundwater hydrology the experiment will be run to estimate parameters for a model, so if the assumption that the model is adequate using correct parameter values can be made, then experiments can be constructed that minimize the prediction error of the model if we knew the structure of the errors. In this way experiments are run on the model to observe its behavior, and various error structures can be assumed to mimic the actual system.

One way of comparing designs is to measure the parameter estimates' covariance structure, or the information structure. Designs constructed this way are called D-optimal if the measure is the determinant of the covariance matrix an it is minimized, A-optimal if the trace of the covariance matrix is used and it is minimized, and a host of others. These have various computational difficulties which leads one to want to find a simpler criterion that still makes sense.

When constructing the covariance matrix or the information matrix, the model sensitivities appear in the formulation as the Jacobian matrix (of the model). The inverse of the covariance matrix is the information matrix. The Jacobian gives the transformation that describes how errors will propagate through the estimation scheme for a particular design. A natural approach is to sample at locations where the model is very sensitive to changes in the parameters, and in this way hopefully, obtain good estimates. This is equivalent to maximizing information (i.e. some measure of the estimates information matrix). If the measure is the determinant then the design is again called D-optimal. If the measure is the trace, it is equivalent to the sum of squared sensitivities. This computationally simple measure (trace) is used in this research.

The second step is a technique to choose designs so that the performance criterion is minimized. In D-optimal or A-optimal design over discrete space the design selection requires a combinatorial approach. In continuous space, at least for groundwater problems, it was found that smooth minimization did not perform well. Using the trace for the information matrix in discrete space again leads to a combinatorial approach, but the ease of computation of the performance criterion makes the problem more tractable.

The third step is to consider design robustness. A design should perform well for a range of parameter values. Additionally the marginal decrease in information of a slightly sub-optimal design should be small.

This research uses a numerical solution to the advection dispersion equation for modelling mass transport in the saturated groundwater domain. Unsteady hydraulics are modelled. These models are used in conjunction with the trace of the information matrix to design a sampling network that maximizes information while satisfying a budget constraint. The resulting designs are intuitivly reasonable. Design of a network configuration as well as a sample schedule jointly is discussed although not solved. Using a Monte Carlo approach it was found that the criterion used leads to designs that are optimal for a narrow parameter range, and more alarmingly that the marginal decrease in information by using a suboptimal design is large.

2. LITERATURE REVIEW

The general topic of experimental design is an active area of research. Researchers in statistics, biostatistics, automatic control, physics and other areas frequently apply experimental design techniques. The topic is introduced in many introductory statistics texts with regard to linear regression for location of the prediction variables (Mendenhall et al., 1986). Another example described by Stein (1945) is an adaptative sampling approach for hypothesis testing where the design variable is sample size. Applications to groundwater resources have been investigated by several researchers in the design of optimal pumping tests (Yeh and Sun, 1984; Hsu, 1984; McCarthy, 1988; Nishikawa, 1988; Hsu and Yeh, 1989).

Steinberg and Hunter (1984) give a review and history of experimental design problems and solution schemes. They elaborate on the importance of sequential designs, where simple designs are used first to locate regions that promise high performance, then detailed design are used to explore these regions. In problems were budget restricts sampling to two rounds (reconnaissance level and intensive level), this approach may not be practical. They point out that optimal design strategies are complicated by the fact that experiments must be planned in the face of considerable uncertainty, and algorithms must often be chosen for desirable combinatorial properties. Finally they describe the difference between model robust designs which give reasonable results for the proposed model even though it is known to be inexact, and model sensitive designs which facilitate improvement of the proposed model by trying to highlight suspected inadequacies. This illustrates the multiobjective nature of experimental design. In a geophysical problem of engineering importance, there may not be time to refine a model, and a robust approach is preferred so the project of interest can proceed.

In the field of automatic control, the experimental design of a lumped parameter system of ordinary differential equations is well developed (Astrom and Eykohoff, 1971; Federer and Balaam, 1973; Mehra, 1974; St. John and Draper, 1975; Ash and Hedayat, 1978). These papers allocate the total number of observations among all possible observation locations with the cost of any observation uniform at any space and time point. In a practical geophysical problem costs vary with location, and it may be preferable to ignore a very informative but costly location in order to meet an experimental budget.

Qureshi (1980) applied D-optimality to identify the optimal location of sensors for two problems: a heat diffusion process and a vibrating string. In the first case it was found that periodic boundary perturbations eliminated computational difficulties. This suggests that the natural frequency of the system dictates in behavior and information on the operating frequencies gives complete knowledge of the behavior of the system. In this spirit Rafajlowicz (1981) suggested estimating the natural frequencies of the eigenvalues of the system rather than the parameters directly and seeking an optimal allocation of sensors over a set of spatial points. In a later paper (1982) an optimal experimental design problem was considered in which both spatial and temporal characteristics of an input signal were optimized from the point view of estimation accuracy expressed in terms of the determinant of the information matrix.

St. John and Draper (1979) reviewed the D-optimal design criterion and discussed algorithms for constructing such designs. In a D-optimal design the covariance matrix for the estimated parameters must be constructed for each candidate design. The algorithms are essentially allocation algorithms, with various schemes for updating the covariance matrix.

Moody and Maddock (1972) suggest how to use a regional planning model to allocate resources for collection for further information. The model is a constrained minimum cost model and the concept of opportunity loss is used to evaluate data quality. The candidate system configuration is already known, as such the problem is one of scheduling and not configuration. The solution is given in a comparison paper (1972) by multiobjective sensitivity analysis.

Hsu (1984) formulated an experimental design problem of a groundwater flow system. His objective was to minimize the cost of a pumping test subject to a set of constraints describing candidate configurations, schedules, and required estimate reliabilities. The reliability measure used was the trace of the covariance matrix, or A-optimality. A heuristic searching algorithm was used to approximate solutions to the resulting mixed integer programming problem.

Yeh and Sun (1984) proposed a systematic procedure for choosing an optimal pumping test whose data would be used to collaborate and verify a simulation model that would be acceptable for the management objective of the aquifer under study. The performance criterion was termed δ-identifiability. A pumping test is δ-identifiable if it produces sufficient data to guarantee that the parameter estimates of the simulation model will yield predictions that are sufficiently accurate for the overall management objective.

This approach is a logical extension of work by Kuszta and Sinha (1978). In their work an experimental design for a distributed parameter system was conducted in which an optimal input signal was selected by maximizing the difference between system outputs observed for two prescribed sets of parameters. In this manner one constructs a most powerful discrimination test to tell which parameter set is in force using the optimal excitation for the test.

McCarthy (1988) obtained solutions using the concept of δ-identifiability for the pumping test design problem for a hypothetical aquifer where the uncertain parameter was aquifer transmissivity. The objective was to minimize the cost of the test subject to configuration constraints, while ensuring the resulting estimates satisfy the δ-identifiability requirement. The procedure used a heuristic searching algorithm.

Nishikawa (1988) used the concept of D-optimality to generate optimal pumping tests for a hypothetical aquifer where the uncertain parameter was transmissivity. The objective was to minimize the cost of the test subject to configuration, schedule, and reliability constraints. The procedure used a global searching algorithm.

Kaunas and Haimes (1985) apply risk analysis concepts to the management of a groundwater contamination problem. The approach is multiobjective and uses tradeoff analysis to evaluate various strategies. Parameter uncertainity is quantified by a sensitivity index where tradeoff among various factors can be evaluated.

Neuman (1980) studied the influence on the variance of parameter estimates to the acquisition of additional data. He illustrates that variance reduction is obtained with additional data and gives some perspective to the quality requirements of these data. The measure used to quantify the variance reduction is the trace of the covariance matrix of the estimates.

Carrera et al. (1984) combine kriging and branch and bound methods to locate optimal observation sites for sampling of fluoride concentrations in an aquifer. The candidate sites are already located and installed so that the problem is one of selection from existing capacity. The performance criterion used is minimization of the estimation variance of the average fluoride concentration. They do not describe how to choose a minimal number of sites, instead they parameterize on the total number of sites to use in a design, which is certainly a good surrogate for a cost constrained strategy. However, unless the individual samples are very expensive it would be foolish not to use all the existing sites since they are already installed.

Wagner and Gorelick (1987) studied statistical implications of transport parameter identifications. They conclude that nonlinear regression provides reliable estimates of average pore water velocity using either temporally or spatially distributed sample points. Using column configurations they found that dispersion, decay, and production coefficients are estimated with significantly more reliability using spatially distributed sample points, and estimates in general are significantly more reliable if sampling is distributed in both space and time. They did not investigate inhomogeneous cases, however they stated that such studies might be useful.

Knoppman and Voss (1987) studied the behavior of sensitivities in one-dimensional solute transport equations and the implications for parameter estimation and sampling design. They found that parameters are most accurately estimated at points with high sensitivity to the parameter, but designs which minimize the variance of one parameter may not simultaneously minimize the variance of other parameters. They reported that maximum sensitivity to velocity is a function of spatial location and hence experimental duration, while maximum sensitivity to dispersion is a function of sample frequency. This reinforces the desirability of sampling at points distributed in space and time. Finally they studied the effect of various experimental designs on the determinant of the estimated parameter's covariance matrix, and found that the design with the smallest determinant tended to give the most reliable estimates with regard to estimate variance.

Strecker et al., (1985) studied the data requirements for transport modelling using the USGS MOC model. A residual analysis was performed to evaluate six sampling strategies, choosing the best as that with the smallest residual as predicted concentrations. They found it was preferable to add new observations in time at existing locations rather than new locations, but it was also observed that the incremental improvement diminished over time.

Myer and Brill (1988) used a facility location model to select monitoring sites to maximize plume detection in a contamination problem under parameter uncertainty. Their approach is a Monte Carlo analysis where many plumes are generated and a binary integer program is used to maximize the likelihood of detection. In their model the cost of all design points are equal.

3. OBJECTIVES

This research presents a systematic procedure for the design of sampling networks enabling assessment of contaminant migration behavior in groundwater systems. The design variables are the location of sample points (configuration) and the timing of sampling schedule. The performance criterion is a weighted sum of squares of model sensitivities. Under certain assumptions it is equivalent to the trace of the estimates information matrix (differing by a constant). Candidate designs are cost constrained, and tradeoff curves are generated to give insight of how sampling resources should be allocated between configuration and schedule. Risk structure is incorporated directly into the sensitivities to investigate its effect on the design selection. These investigations require;

1. A simulation model that describes hydraulics and mass transport. For this study, a simple finite difference model is used;

2. A method to calculate sensitivities of the model to changes in the transport parameters. These are used to construct the performance criterion used to select a design. The sensitivities are calculated numerically using the influence coefficient method;

3. A performance criterion used to evaluate candidate designs;

4. A selection algorithm which optimizes the performance criterion over the design variables subject to economic constraints.

Each of these tools are discussed in the following sections.

4. SIMULATION MODEL

The design problem should be able to incorporate any simulation model. For development, this research uses a simulation model in one and two dimensions. The two dimensional formulation discussed below, consists of decoupled flow and transport equations.

Two-dimensional, heterogeneous, isotropic confined flow in porous media is governed by the following equation (Bear, 1972, 1979):

$$S \frac{\partial H}{\partial t} = \frac{\partial}{\partial x} \left[K_{xx} \frac{\partial H}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_{yy} \frac{\partial H}{\partial y} \right] + M \tag{1}$$

subject to the following initial and boundary conditions:

$$H(x,y,0) = \text{known} \quad (x,y) \in \Omega \tag{2}$$

$$H(x,y,t) = \text{known} \quad (x,y) \in d\Omega_1 \tag{3}$$

$$\left[K_{xx} \frac{\partial H}{\partial x} \right] \frac{\partial x}{\partial n} + \left[K_{yy} \frac{\partial H}{\partial y} \right] \frac{\partial y}{\partial n} = \text{known} \quad (x,y) \in d\Omega_2 \tag{4}$$

where

H	=	hydraulic head (L)
K_{xx}	=	hydraulic conductivity (x plane, x direction, L/T)
K_{yy}	=	hydraulic conductivity (y plane, x direction, L/T)
S	=	storage coefficient
M	=	net injection/extraction rate (L/T)
Ω	=	flow region
$d\Omega$	=	boundary of flow region ($d\Omega_1 \cup d\Omega_2 = d\Omega$)
$\frac{\partial}{\partial n}$	=	normal derivative to boundary

The governing equation of the solute transport process used in this study is given by (Bear, 1972, 1979);

$$R\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}\left[D_{xx}\frac{\partial C}{\partial x} + D_{xy}\frac{\partial C}{\partial y}\right] + \frac{\partial}{\partial y}\left[D_{xy}\frac{\partial C}{\partial x} + D_{yy}\frac{\partial C}{\partial y}\right]$$

$$- V_x\frac{\partial C}{\partial x} - V_y\frac{\partial C}{\partial y} + M \tag{5}$$

subject to the following initial and boundary conditions:

$$C(x,y,0) = \text{known} \qquad (x,y) \in \Omega \tag{6}$$

$$C(x,y,t) = \text{known} \qquad (x,y) \in d\Omega_1 \tag{7}$$

$$\left[C\frac{V_x}{R} - \frac{D_{xx}}{R}\frac{\partial C}{\partial x} - \frac{D_{xy}}{R}\frac{\partial C}{\partial y}\right]\frac{\partial x}{\partial n} = \text{known} \qquad (x,y) \in d\Omega_2 \tag{8}$$

$$\left[C\frac{V_y}{R} - \frac{D_{yx}}{R}\frac{\partial C}{\partial x} - \frac{D_{yy}}{R}\frac{\partial C}{\partial y}\right]\frac{\partial y}{\partial n}$$

where;

C	=	mass of solute per volume of medium (M/L^3)
D_{xx}, D_{xy} ...	=	components of hydrodynamic dispersion tensor
R	=	retardation factor
V_x	=	average fluid velocity in x direction (L/T)
V_y	=	average fluid velocity in y direction (L/T)
M	=	net mass injection/extraction rate (M/T)
Ω	=	flow region
$d\Omega$	=	boundary of flow region ($d\Omega_1 \cup d\Omega_2 = d\Omega$)

The above equation reflects certain implicit assumptions. First there is no generation or decay of the solute. Also, adsorption is described by a linear equilibrium isotherm, hence the use of a retardation factor (Bear and Verruijt, 1987). Included are terms for sources and sinks. Imbedded in all terms is the porosity which permits the transformation of mass per fluid volume and mass per medium volume. The hydrodynamic dispersion coefficients for an isotropic porous medium are expressed by (Bear, 1979);

$$D_{xx} = (\alpha_L - \alpha_T)\,\frac{V_x^2}{V} + \alpha_T V + D^* \tag{9}$$

$$D_{xy} = (\alpha_L - \alpha_T)\,\frac{V_x V_y}{V} \tag{10}$$

$$D_{yy} = (\alpha_L - \alpha_T)\,\frac{V_y^2}{V} + \alpha_T V + D^* \tag{11}$$

where

V	=	$(V_x^2 + V_y^2)^{1/2}$
α_L	=	longitudinal dispersivity
α_T	=	transverse dispersivity
D^*	=	molecular diffusivity

The distribution of average velocities is computed using Darcy's Law, written as

$$V_x = \frac{K_{xx}}{n}\,\frac{\partial H}{\partial x} \tag{12}$$

$$V_y = \frac{K_{yy}}{n}\,\frac{\partial H}{\partial y} \tag{13}$$

where n is the average porosity of the porous medium.

The transport equation is coupled to the flow equation through the average velocities. If the solute significantly alters the density of the solvent (water) as a function of concentration, then the flow model must be altered; the two equations coupled and solved simultaneously. For simplicity it is assumed that this is not the case, and the flow equation can be solved independently.

Various finite difference, finite element and other methods have been proposed for the numerical solution of these partial differential equations. For simplicity an explicit finite difference scheme is employed to solve these equations. An upwind formulation is used for the transport equation to alleviate overshoot and undershoot associated with the numerical solution when advection dominates.

Solutions to equations (1) and (5) were obtained using a forward Euler scheme. Centered difference approximations were used for the spatial discritization equation (1). The stability criterion for the flow model is the smaller of;

$$\Delta t \leq \frac{1}{2} K_{xx} \Delta x^2$$

$$\Delta t \leq \frac{1}{2} K_{yy} \Delta y^2$$

where

Δx	=	characteristic length of discritized domain in x direction
Δy	=	characteristic length of discritized domain in y direction

In equation (5) the space discritization of the second order partials is the same leading to an analogous stability criterion with D_{xx}, D_{yy} replacing the conductivities above. The first order partials are approximated using forward or backward differences depending on the local velocity direction (upwind formulation) when the local Peclet number exceeds 2.0. The Peclet number in the x direction is computed from;

$$P_{e_x} = \frac{|V_x| \Delta x}{||D||}$$

where

$||D||$ is some norm of the dispersion tensor.

In this study $||D|| = (D_{xx}^2 + D_{yy}^2)^{1/2}$.

When velocity is significant an additional stability criterion must be met; (Courant criterion)

$$\Delta t \leq \frac{\Delta x}{V}$$

where

V = average velocity

For this study stability was determined by trial and error, Δt being adjusted until the scheme remained stable for the duration of the simulation.

5. SENSITIVITY COEFFICIENTS

The sensitivities are used to construct an information measure that evaluates the information available in a particular design. Sensitivities in this research are defined to be the partial derivative of the observed state variable (concentration) with respect to a particular parameter. The sensitivity will change over time at a particular spatial location, as well as change over space for a particular time, thus the sensitivities are time and space dependent functions.

Three methods are available to compute the sensitivities of the solution to changes in the parameters; the influence coefficient method, the sensitivity equation method, and the variational method (Yeh, 1986). In this study the influence coefficient method is used. In this method one perturbs the parameter settings and solves the simulation model again. These solutions along with the original are used to form difference quotients which are the sensitivity coefficients. The method must be used with some caution as its accuracy is significantly affected by the size of the perturbation. The influence coefficient method requires the least amount of computer storage, and can produce accurate sensitivities (Wang, 1987). Once the sensitivities are computed for each parameter at each point in space and time they are assembled into the familiar Jacobian matrix. For this study sensitivity of head to variations in hydraulic parameters is not considered, and only sensitivity of concentration is considered.

The sensitivity at some point in space and time is computed from:

$$\frac{\partial C}{\partial \theta_j}(x,y,t,\theta) = \frac{C(x,y,t,\theta + \Delta_j) - C(x,y,t,\theta)}{\Delta_j} \tag{14}$$

where

$$\theta + \Delta_j = (\theta_1, \theta_2, \cdots \theta_j + \Delta_j, \cdots \theta_k)^T$$

These sensitivities are then assembled into the Jacobian matrix, for instance suppose a two point design is used (i.e. $\{(x_1, t_1); (x_2, t_2)\}$) then the Jacobian would have the form

$$J = \begin{bmatrix} \frac{\partial C}{\partial \theta_1}(x_1, t_1, \theta) & \frac{\partial C}{\partial \theta_2}(x_1, t_1, \theta) & \cdots & \frac{\partial C}{\partial \theta_k}(x_1, t_1, \theta) \\ \frac{\partial C}{\partial \theta_2}(x_2, t_2, \theta) & \frac{\partial C}{\partial \theta_2}(x_2, t_2, \theta) & \cdots & \frac{\partial C}{\partial \theta_k}(x_2, t_2, \theta) \end{bmatrix} \tag{15}$$

6. PERFORMANCE CRITERION

The performance criterion for this study is a function used to evaluate a particular design. The criterion uses the sensitivities to measure information in a particular design. Other popular criteria also use the sensitivities in the form of approximations of the estimate's covariance matrix. As compared to other criteria, this criterion is computationally simple and intuitive.

The criterion uses the concept that parameters are most accurately estimated at points with high sensitivity to the parameter (Knoppman and Voss, 1987). An intuitive criterion is then one that sums up the values of sensitivities at some point in space in time. However, sensitivities can be either positive or negative, so to prevent cancellation of high but opposite sign sensitivities the squared sensitivity is used. Weights are included to reflect relative importance of parameters. The weighted sum squares criterion at some sample point i is written as;

$$Z_i = \sum_{j=1}^{k} \omega_j \left[\frac{\partial C_i}{\partial \theta_j}\right]^2 = t_r \left(J_i^T J_i W\right) \tag{16}$$

where

Z_i	=	information at ith location
J_i	=	Jacobian matrix at ith location
ω_i	=	weight on jth parameter
W	=	diagonal matrix of weights which has determinant one
k	=	number of parameters

Allowing W to be any non-negative matrix generates different criteria, some which may include cross sensitivities. Here only diagonal matrices are used which is a special class of the more general criteria discussed by Sacks and Ylvisaker (1968). **For this study,** the sum of squares criterion is called the CR-optimal criterion.

Other popular criteria are A-optimality and D-optimality. These operate on the estimate's covariance matrix which is constructed from the sensitivity information. For an additive error model;

$$C_{Observed} = C_{Model}(\theta) + Error \tag{17}$$

where,

$$E(Error) = 0$$

$$V(Error) = \Sigma$$

Assuming a least squares parameter identification scheme will be used, the linear approximation to the estimates covariance matrix is; (Yeh and Yoon, 1981)

$$V(\theta) = M \, \Sigma \, M^T \tag{18}$$

where

$$M = (J^T J)^{-1} J^T$$

In the special case where

$$\Sigma = \sigma^2 I$$

the approximation of the estimate's covariance matrix is

$$V(\theta) = \sigma^2 (J^T J)^{-1} \tag{19}$$

The A-optimality measure is the trace of the covariance matrix, while the D-optimality measure is the determinant of the covariance matrix. The information matrix for the special case above is

$$I(\theta) = \frac{1}{\sigma^2} (J^T J) \tag{20}$$

Using these measures the formal design problem can be stated as;

D-optimality

$$\begin{aligned} \min \quad & \det\left(V_z(\theta)\right) \\ \text{s.t.} \quad & \text{cost}\,(z) \le \text{budget} \\ & z \in Z \end{aligned} \tag{21}$$

A-optimality

$$\begin{aligned} \min \quad & \operatorname{tr}\left(V_z(\theta)\right) \\ \text{s.t.} \quad & \text{cost}\,(z) \le \text{budget} \\ & z \in Z \end{aligned} \tag{22}$$

CR-optimality

$$\begin{aligned} \max \quad & \operatorname{tr}\left(I_z(\theta)\right) \\ \text{s.t.} \quad & \text{cost}\,(z) \le \text{budget} \\ & z \in Z \end{aligned} \tag{23}$$

where

$V_z(\theta)$ = covariance approximation for design z

$I_z(\theta)$ = information approximation for design z

Z = set of all designs

Statistically D-optimality minimizes the volume of the ellipsoid described by the parameter estimate's probability density function, it is also essentially the same as minimizing the Wilk's generalized variance (Mood et al., 1974) hence it is a popular criterion that is mathematically interesting. A-optimality minimizes the volume of the parallelipiped that encloses the ellipsoid for some design (Fedorov, 1972). The weighted sum of squares maximizes the trace of the information matrix for some design. The only part of any of these criteria that changes over design space is J^TJ (or its inverse).

From a computational standpoint the trace of the information matrix is simplest to compute. Sensitivities at every possible design point are computed, squared, weighted and summed to form a "unit" information number for that design point. The total information in a particular design is then computed by summing up the unit numbers for each point in the design. For the A-optimal or the D-optimal criterion, the design must be specified, the Jacobian for that design computed, J^TJ computed and then inverted. In numerical experiments it was found that smooth minimization on analytical solutions did not work well using either A or D optimality. For combinatorial simplicity the sum of squares criterion was chosen at the expense of statistical interpretation, however, the physical interpretation of sampling where sensitivities are high seems to be an acceptable tradeoff.

7. SELECTION ALGORITHM

In this section the optimal design selection is formulated as a finite dimensional integer programming problem. The cost constraints make the problem a fixed cost problem, so the design space is very large. The resulting program is solved by a polynomial time approximation scheme.

Formally let,

nx	=	number of x spatial locations
ny	=	number of y spatial locations
nt	=	number of possible sampling times
S	=	$\{1,2, \cdots ,nx * ny \}$
$c(x_i, y_i)_j$	=	cost of installation of sampler at site (x_i, y_i) and sample collection from time j until nt.
$I(x_i, y_i, t_j, \theta)$	=	unit information number at site (x_i, y_i) and time (t_j) for the parameter set θ

It is assumed that the experiment duration is known. It is also assumed that one sampling has begun at a site, it will be continued until the end of the experiment. This assumption is used to construct the objective function as a sum of time integrated information numbers at each site. For instance, if site (x_i, y_i) is considered, and sampling is begun at time (t_j) and the experiment ends at time (t_{NT}) then the total (time integrated) information available at that point is

$$\text{Total information } (x_i, y_i) = \int_{t_j}^{t_{NT}} I(x_i, y_i, t, \theta)\, dt$$

In the discrete case, the integration is treated as a sum and written as

$$\text{Total information } (x_i, y_i) = \sum_{j=l}^{NT} I(x_i, y_i, t_j, \theta)$$

The assumption is also used in the construction of cost for a site as stated above in the definition.

A zero-one indicator variable is used to identify which sites are selected. A value of one means that site is selected. The zero-one variable is double subscripted, the first subscript is location, the second sample initiation time. For instance

$$z_{2,7} = 1$$

means site 2 (with locations x_2, y_2) with sampling starting at time (t_7) is indicated as a selected point.

The selection problem can be written as;

$$\max_{z_{ij}} \sum_{i \varepsilon S} z_{i1} \left[\sum_{j=1}^{NT} I(x_i, y_i, t_j, \theta) \right] + \cdots + z_{iNT} \left[\sum_{j=NT}^{NT} I(x_i, y_i, t_j, \theta) \right] \tag{24}$$

s.t.

$$\begin{array}{lll} \sum_{i \varepsilon S} z_{i1}\, c(x_i, y_i)_1 + \cdots + z_{iNT}\, c(x_i, y_i)_{NT} & \leq \text{budget} & \\ z_{i1} + \cdots + z_{iNT} & \leq 1 & \forall_i \ \varepsilon\ S \\ z_{i,j} & = 0 \text{ or } 1 & i\ \varepsilon\ S \end{array}$$

The solution of this multidimensional knapsack problem defines the optimal configuration and schedule for the sampling network design. As a simplification, the time integration is done for j = 1,NT only, reducing the problem to that of configuration only. In this form the problem is a 0-1 knapsack problem which is solved by a fully polynomial time approximation scheme as described by Papadimitriou and Steiglitz (1982).

8. COSTS INVOLVED

In designing a sampling network several costs are to be considered. When timing is involved the problem becomes a fixed cost problem over time.

In this study certain sunk costs are already assumed to have been budgeted. The installation, development and operation of the injection and extraction system are assumed known. The costs remaining are sample site installation, sampling and analysis. Sampling and analysis are lumped into a generalized sampling cost which is assumed to be easily computed in practice. Installation cost will be a function of location and depth as well as the cost of the sampler (pump) itself. The ratio of installation to sampling is assumed to be ten. That is it costs ten units of budget to install a site, and one unit per sample after installation.

It is desired of course to install as few sites as possible and take many samples to make the average unit cost per sample small. This is not the only reason, samplers are expensive and the number of these available imposes a practical constraint for any real problem, here only theory is considered but it is important to keep this thought in mind. Marginal information gain has been observed in the literature to decrease over time, and the diminishing return of continued sampling at one installation would possibly justify the addition of another site. This behavior is illustrated in the one dimensional cases to follow. The two dimensional cases ignore sampling costs and consider configuration costs only.

9. ONE DIMENSIONAL EXAMPLE

To illustrate the feasibility of the proposed method, a one dimensional transport scenario is investigated. A steady velocity field is assumed. A pulse of contaminant is injected instantly at some injection site. The volume injected is assumed to be so small that the velocity field is undisturbed. The flow domain is homogeneous and isotropic. Designs are investigated for a system governed by

$$R \frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - V \frac{\partial C}{\partial x} \tag{25}$$

where

R	=	retardation factor
D	=	dispersion coefficient
V	=	average velocity

Concentration is zero at $\pm \infty$, and the pulse is injected at the origin. A physical setup that might be approximated by this model is pictured in Figure 1. The analytical solution for this problem is (Bear (1979):

$$C(x, t) = \frac{M}{\sqrt{4\pi \frac{D}{R} t}} \exp \left\{ -\frac{1}{2} \frac{(x - \frac{V}{R} t)^2}{2 \frac{D}{R} t} \right\} \tag{26}$$

where

M	=	mass injected
t	=	time from injection

Sensitivities are calculated numerically using the influence coefficient method.

Two examples are studied, the first has a dispersion coefficient five times larger than the second. Concentration profiles at a selected times are shown in Figure 2. The first example has the contaminant spreading rapidly, while in the second

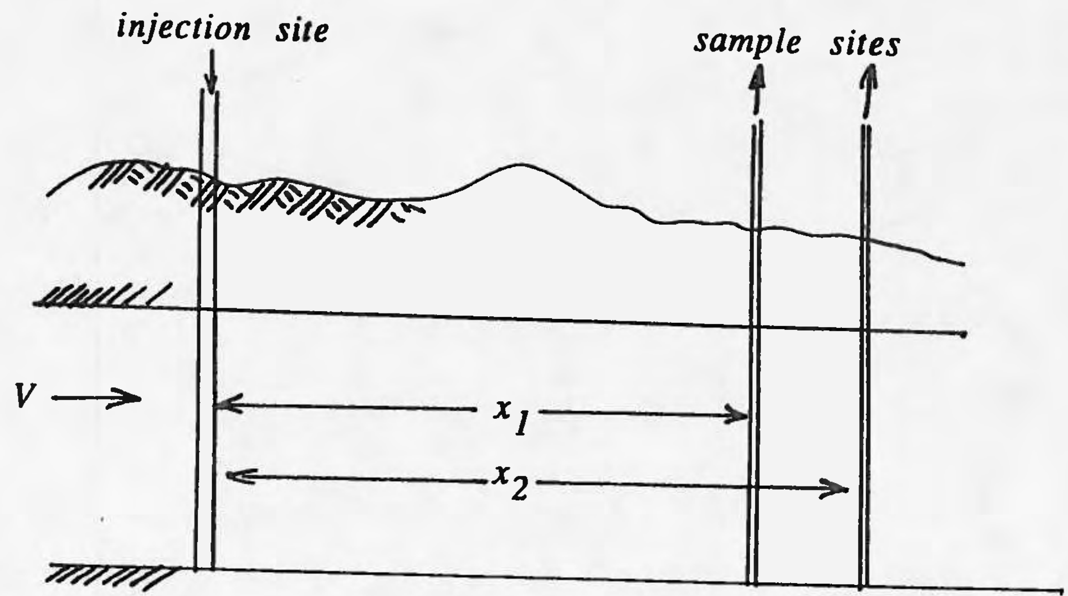

Figure 1 Typical Experiment Configuration

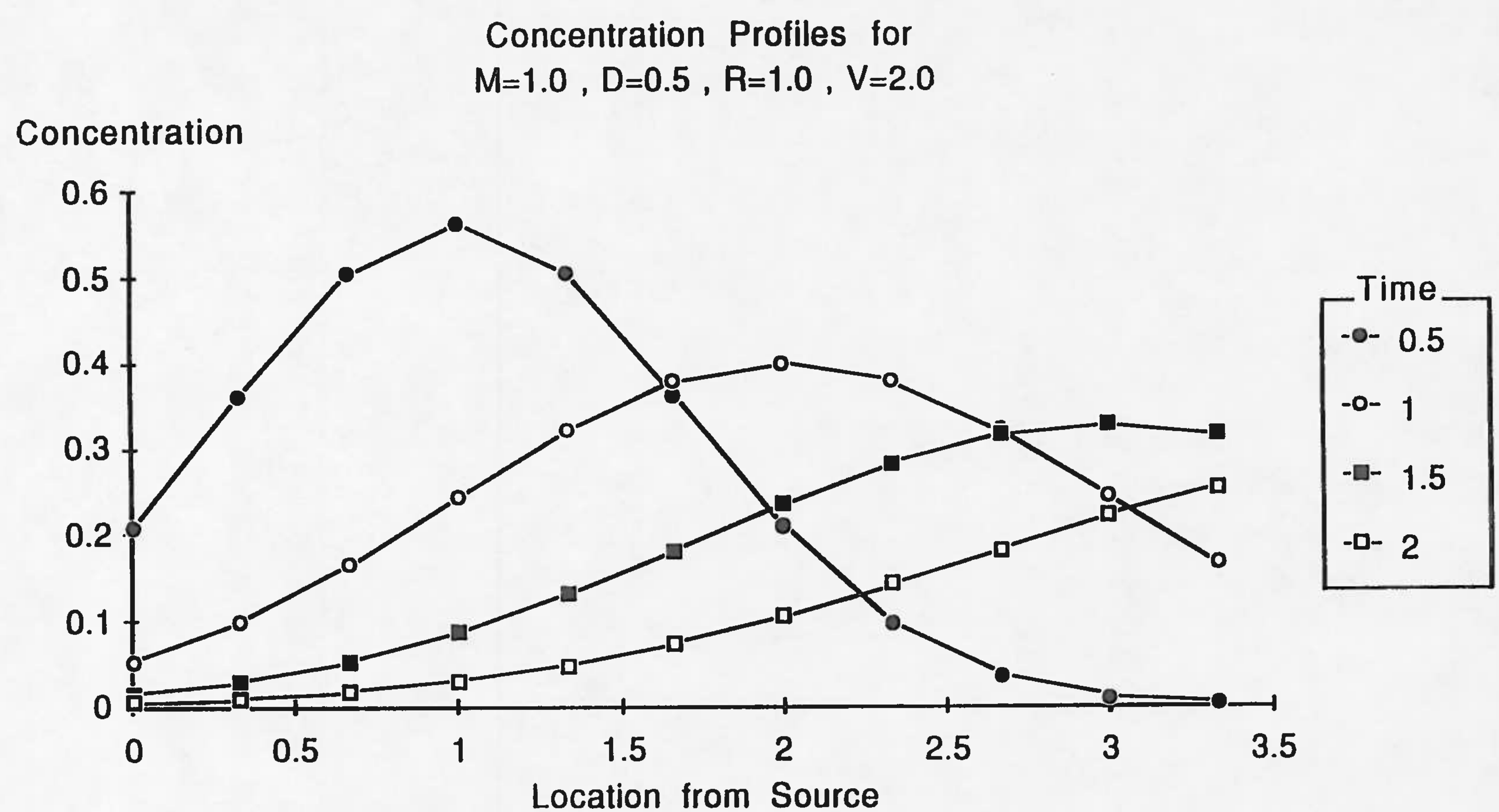

Figure 2 Concentration Profiles for M=1, D=0.5, R=1.0, V=2.0

it moves more as a slug.

The cost of installation of a sample point was assumed to be 1.0 while that of sampling and analysis is 0.1 units. Costs were assumed uniform over space. The design algorithm was constrained so that $x \geq 0.5$ for these examples. If this is not done the injection source is always a maximum information point. In practice a scaling effect is observed where parameter estimates are unstable if sampling is done too near the source early in time.

Figure 3 shows the CR-optimality criterion over time and integrated. We observe that for a single location and single sample the best point is $x = 0.5$, $t = 0.5$. The method instead chooses $x = 2.0$, $t = 0.5$ as the best time integrated single point, when the system is first configured, then the resulting configuration is scheduled. Where the multidimensional knapsack problem solved instead the more obvious answer would have been selected. The problem was decomposed in this manner out of computational convenience. A tradeoff curve for 1 point, 2 point and 3 point designs is shown in Figure 4. The design locations for each of the three curves are $x = 2.0$; $x = 1.5, 2.0$; $x = 1.5, 2.0, 2.5$ respectively.

From Figure 4, the tradeoff curves indicate that within a particular configuration it is profitable to take more samples overall, than to add a location and take less samples. A conclusion is that when spreading occurs quickly it is more advantageous to add sample times rather than add locations.

The concentration profiles for the second example are shown in Figure 5. For this example one can observe that the information moves as a slug in Figure 6. Figure 7 shows the tradeoff curves associated with a 1, 2 and 3 location design. The locations for the 1 point, 2 point and 3 point design are; $x = 1.0$; $x = 1.0, 2.0$; $x =$

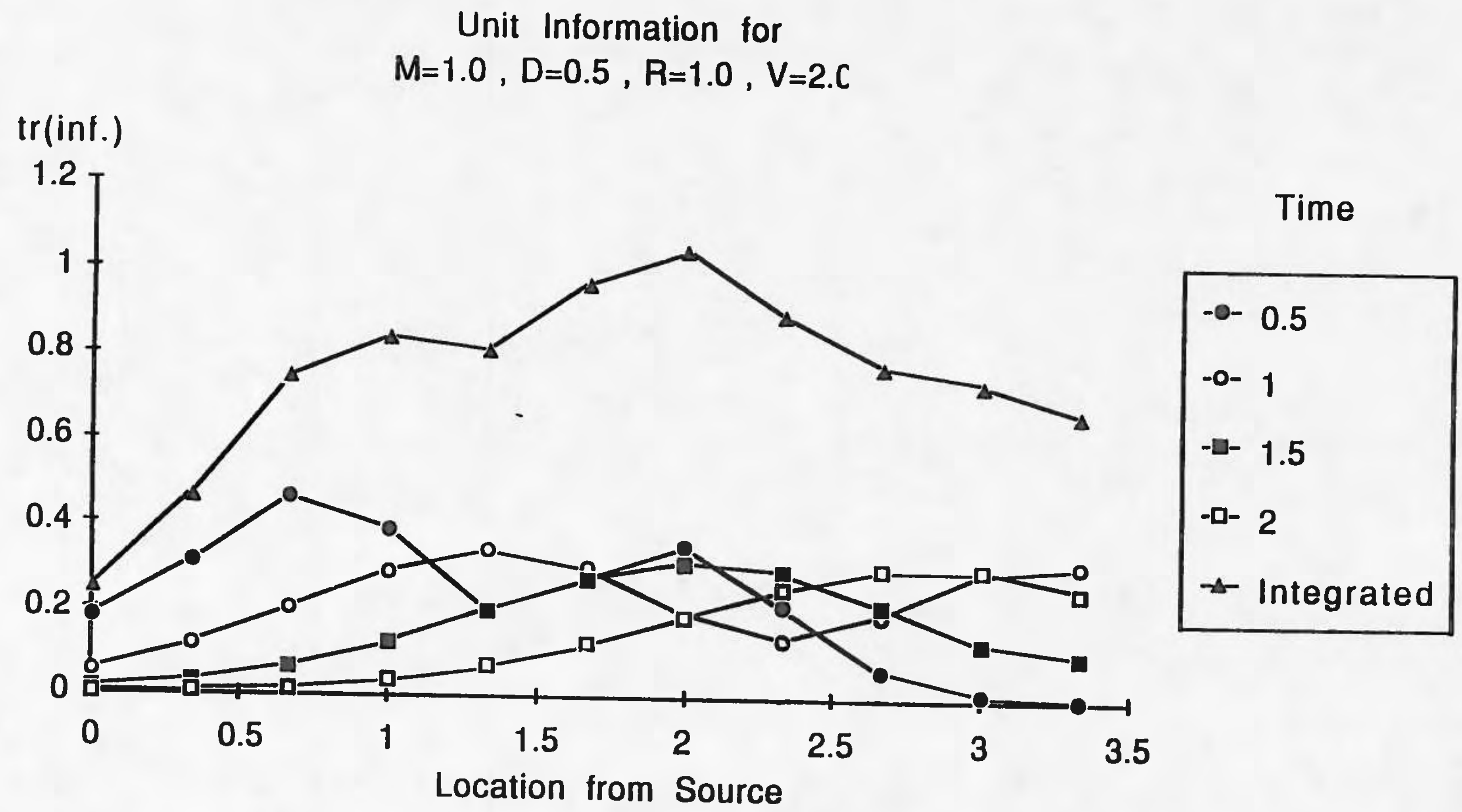

Figure 3 Unit Information for M=1, D=0.5, R=1.0, V=2.0

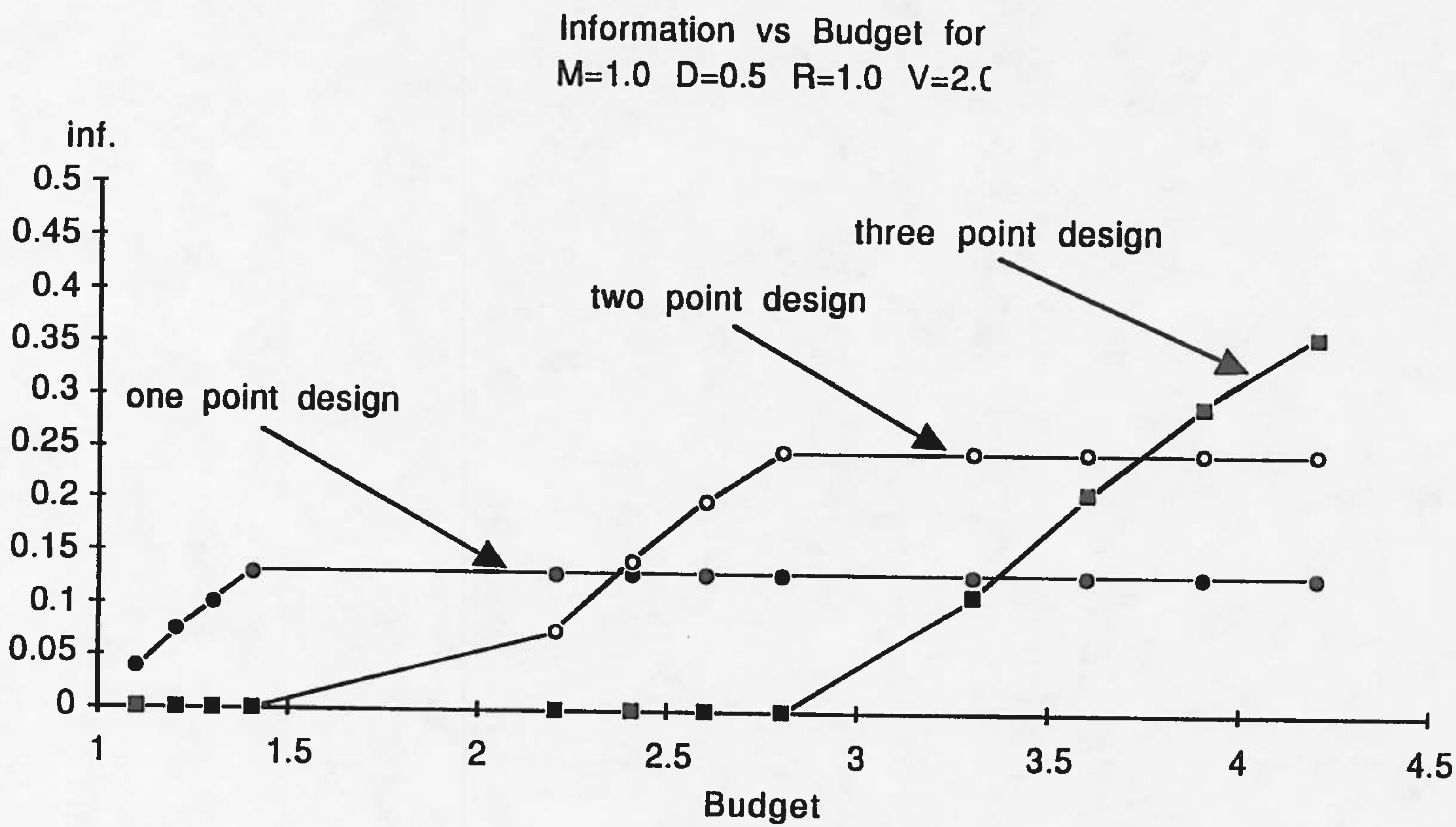

Figure 4 Tradeoff Curve for M=1, D=0.5, R=1.0, V=2.0

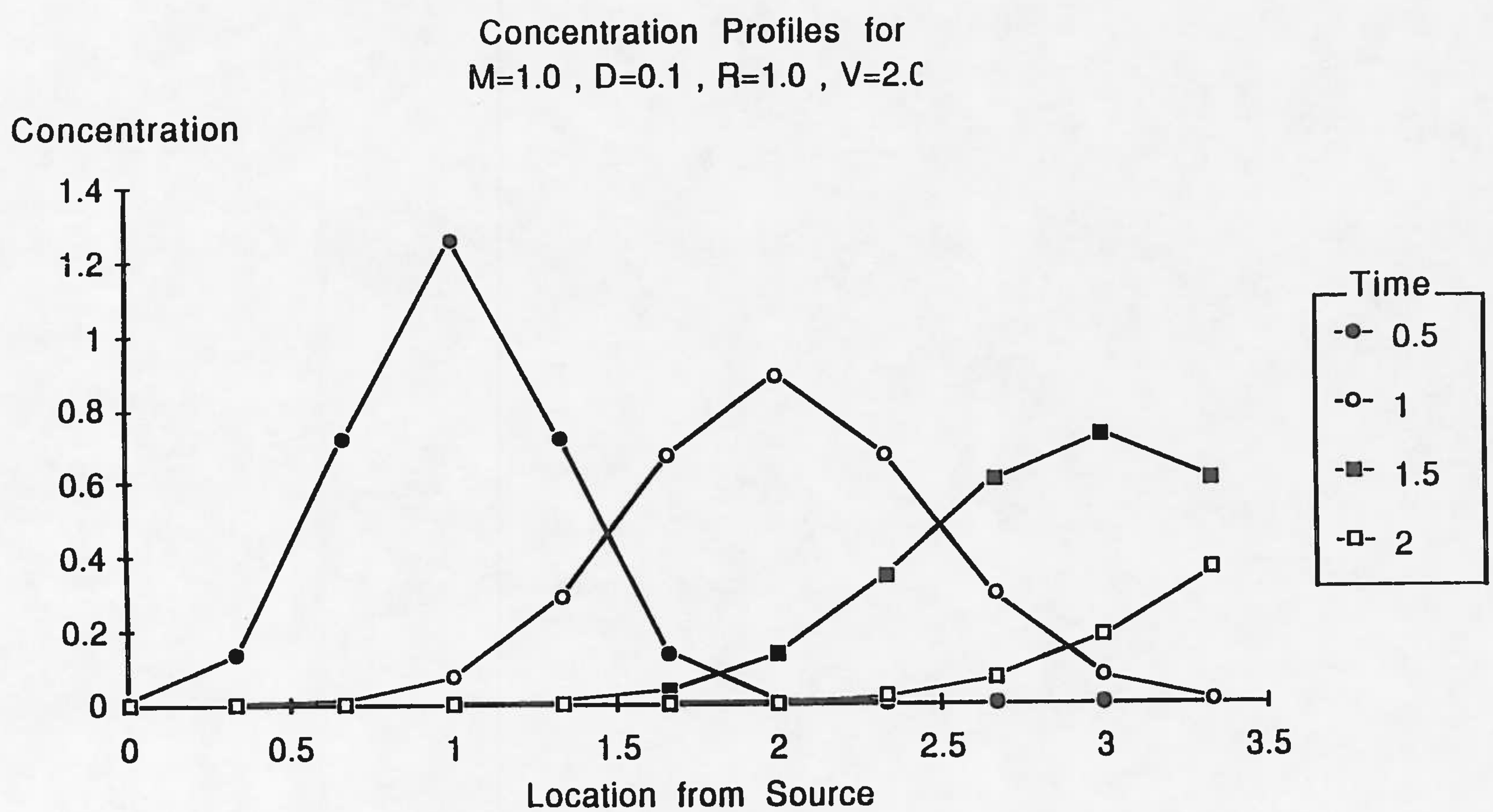

Figure 5 Concentration Profiles for M=1.0, D=0.1, R=1.0, V=2.0

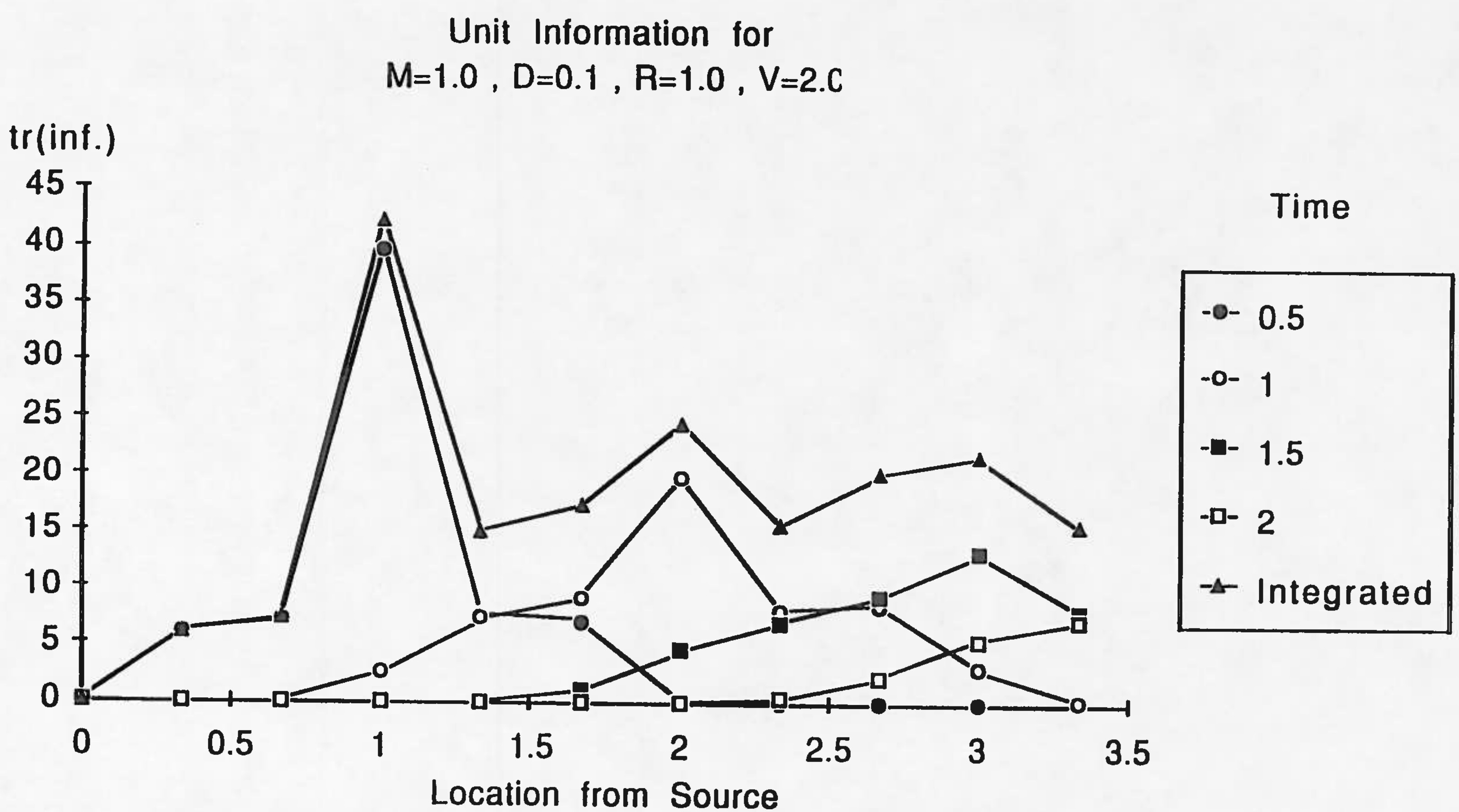

Figure 6 Unit Information for M=1.0, D=0.1, R=1.0, V=2.0

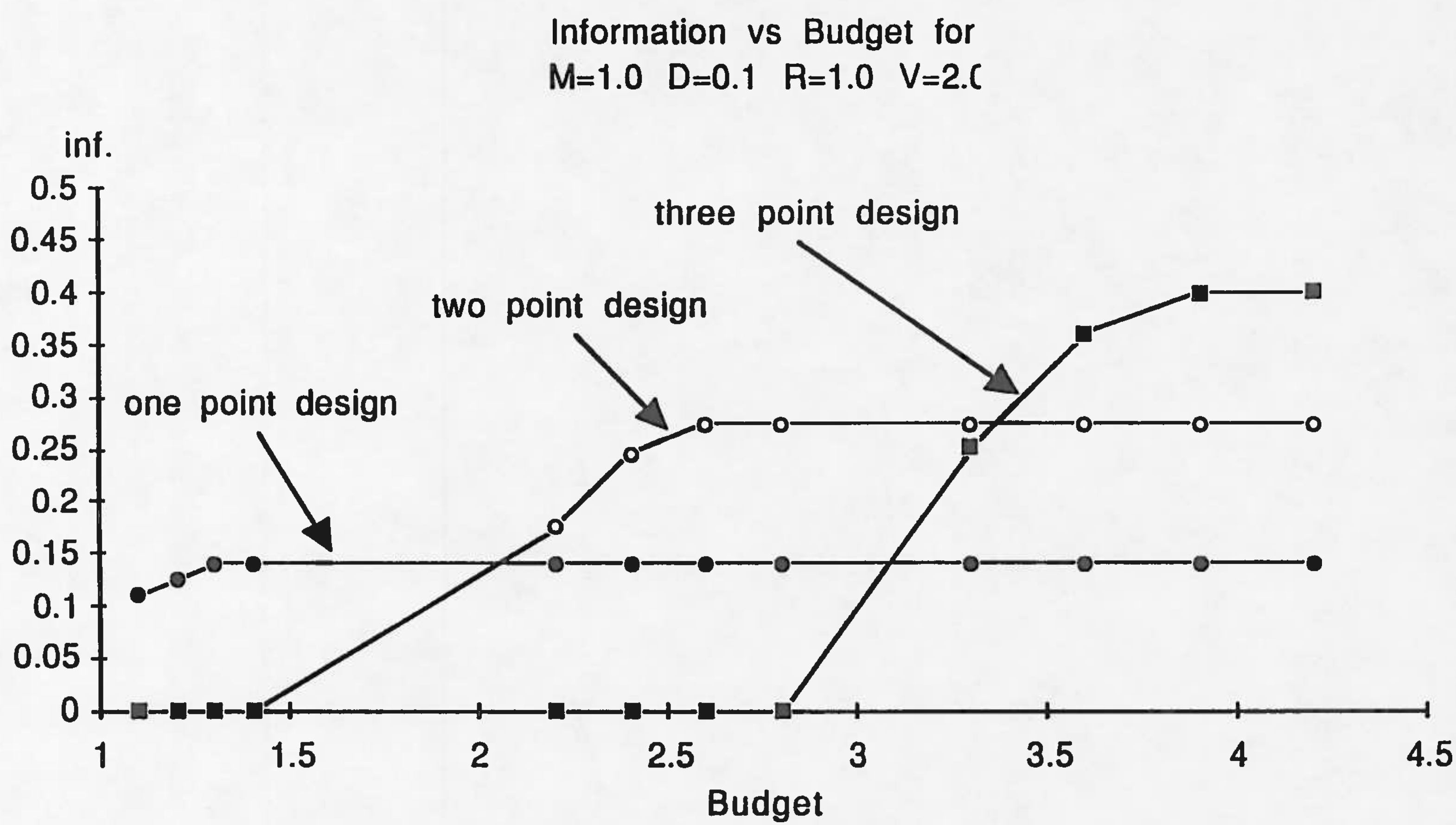

Figure 7 Tradeoff Curve for M=1.0, D=0.1, R=1.0, V=2.0

1.0, 2.0, 3.0 respectively.

From the tradeoff curves a conclusion is that sample location is more significant than overall samples, with information gain being exhausted in each configuration at the higher budgets. This makes sense since the information moves out of the sample region with less spreading. This supports the observation of Strecker et al., (1985) that the incremental improvement of a sampling network as observations are added diminishes over time.

10. TWO DIMENSIONAL EXAMPLES

This section presents some two dimensional examples where the velocity field is unsteady and is computed using equations (1) and (5). The designs are for configuration only, simultaneous configuration and scheduling is not considered.

In all examples, a finite difference grid with a spacing of 0.5 m in both directions is used. The time step used in 0.001 days, with sampling every 0.5 days. The initial conditions are set so that the regional flow is approximately 2.0 m/day when the hydraulic conductivity is 1.0 m/day. The injection mass is 0.001 kg/m^3.

The experiment modeled is run as follows. Injection and extraction are started at time 0.0 days. Mass is added at injection sites for 2.0 days and then stopped, but the hydraulic stress is left on for the entire experiment. The modeled duration is 10.0 days, but sampling is not initiated until all the mass has been added. This hypothetical experiment is chosen because of its simplicity but also because it is somewhat representative of some field experiments.

Cost of sampling is assumed uniform in all cases so the only effect of changing budget will be to choose the number of design locations in an optimal design.

10.1 Case 1

Case 1 is a homogeneous case with a doublet aligned 30° off the regional flow direction. Figure 8 shows the solution domain used. Table 1 shows the parameters used for Case 1.

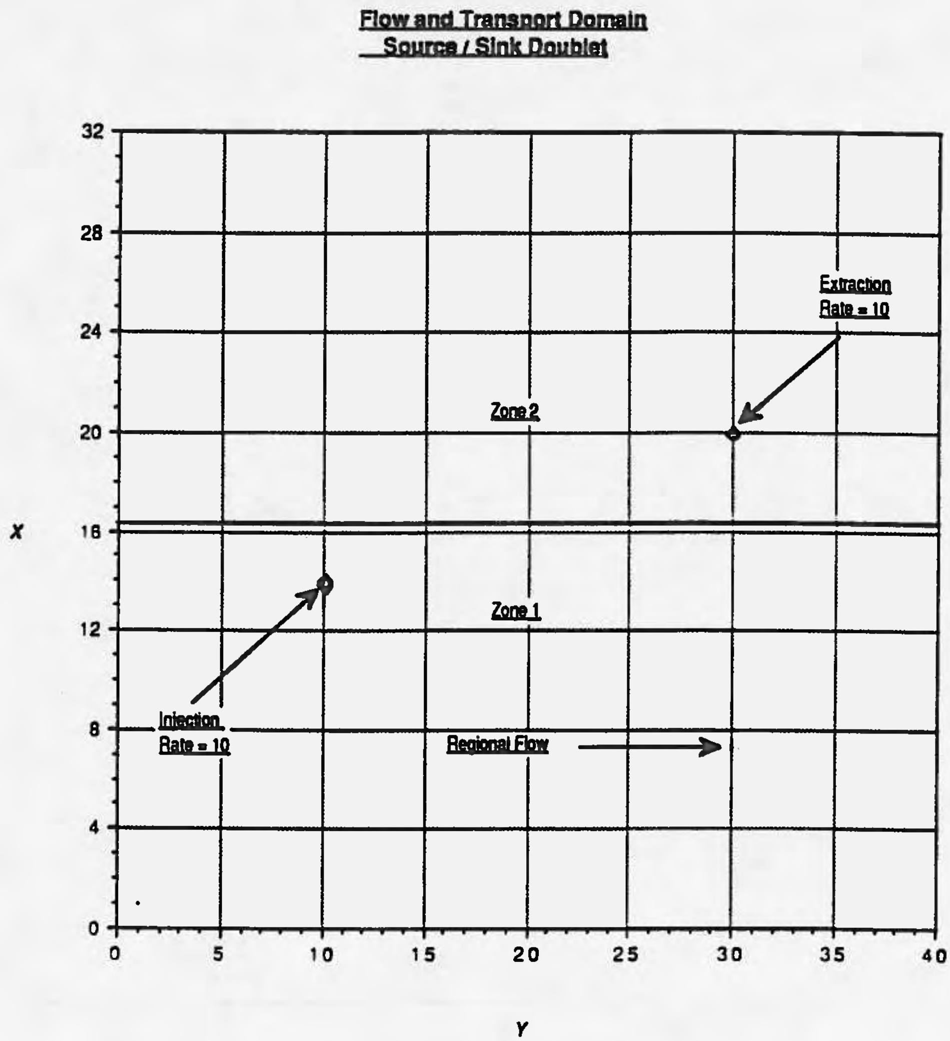

Figure 8 Flow and Transport Domain (Cases 1-7)

	Zone 1	Zone 2	Weight
K_x	3.0	3.0	1.0
K_y	3.0	3.0	1.0
S	0.1	0.1	1.0
α_L	0.5	0.5	1.0
α_T	0.5	0.5	1.0
R	1.0	1.0	1.0

Table 1
Parameters for Case 1

The column for weights gives the weights assigned to each parameter in computing the objective function.

Figure 9 shows the CR-objective surface for Case 1. The maximum information is located at the injection site. In nearly all runs this was observed to be true. In retrospect this should have been expected based on the selection criterion, consider how the sensitivities are formed.

A parameter set is specified and a plume of mass is generated. Then to compute sensitivities the parameters are perturbed and other plumes are generated. With respect to the first parameter set, the subsequent plumes differ from the first only in shape and location. The sample location where a change in plume location would be most sensitive would surely be the injection site or nearby for so short duration an experiment. At this sample location mass is concentrated early in time and small parameter changes will cause large changes in the location of the centroid of the mass distribution, and hence large sensitivities.

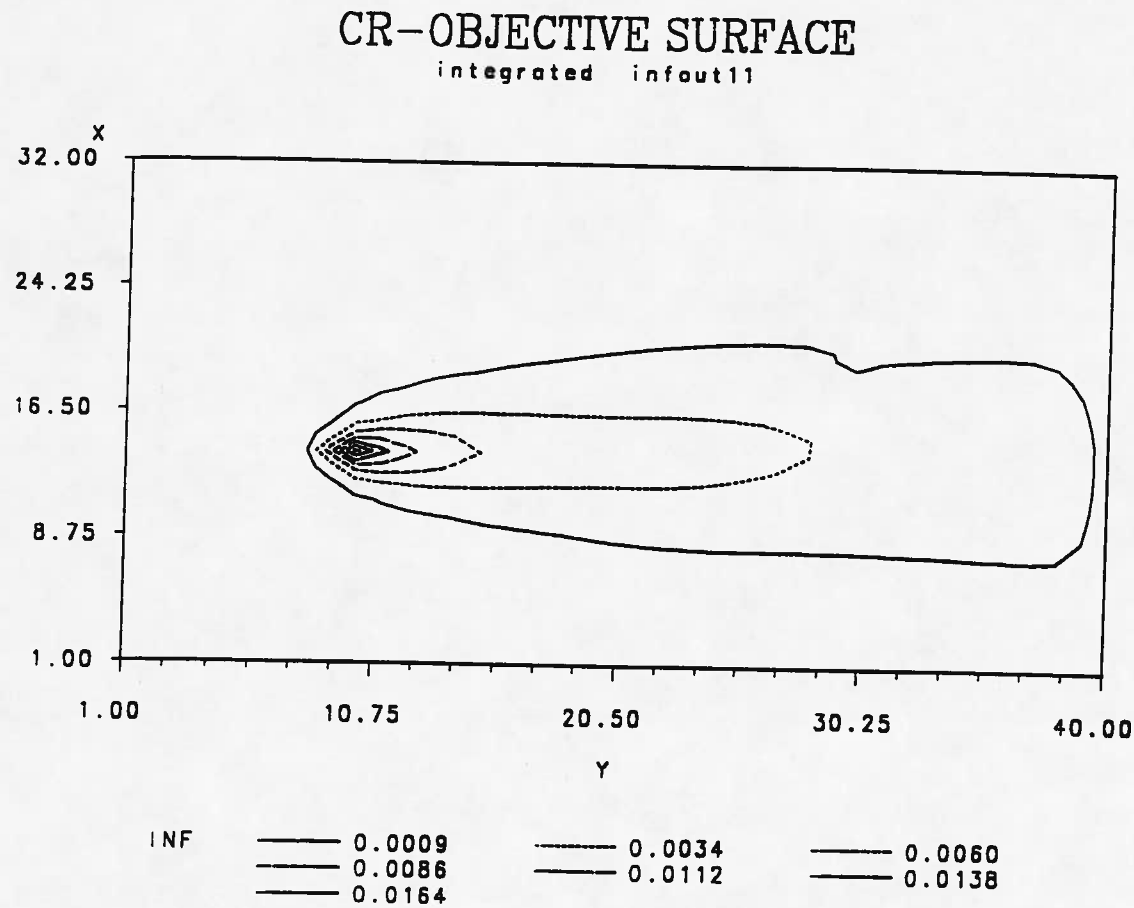

Figure 9 Objective Surface Case 1

It is unfortunate that this occurs, because although it makes complete sense mathematically and conceptually, it leads to rather uninteresting designs. Furthermore the designs are unrealistic because the parameters of interest should represent regional values (averages) relative to the scale of the problem of interest. Sampling so close to the injection site will give at best a local estimate and will introduce the question of scaling effect that has been observed in practice.

The "scaling effect" which has been mentioned is the name given to the observation that dispersivities increase with distance between the source and observation point. It is postulated that in a well behaved system the dispersivities will reach an asymptotic value and the classical advection-dispersion equation used here would be appropriate.

Using a stochastic approach to model transport behavior Dagan (1982) has demonstrated asymptotic behavior. For the classical equation the implication is not to sample too near a source early in time.

A physical interpretation is offered by Naff et al. (1988). An injected slug will preferentially displace the ambient fluid in the vicinity of a well where the medium is more conductive. In a vertically averaged sample, concentration will not be uniform at the scale of interest. At some distance from the source mixing (averaging) would take place and classical behavior will be observed. The implication is that early sampling will not allow for adequate mixing.

A rule of thumb for use of the "classical" equation is that the averaging scale is on the order of 10 times the average characteristic length of the geologic structure of the medium. In some cases this can be hundreds of meters. (Marsily, 1986).

In terms of the designs presented here the scaling effect is a vestige of the underlying model used and not the approach itself. When this is considered the apparently uninteresting designs offered make good sense mathematically (with respect to the selection criterion and model).

For the time being, this is ignored since the results here are still useful, not for actual design, but to highlight some considerations that may be used when considering a design.

Figure 10 shows the optimal six point design. The design is reasonable if the near source problem is ignored. A suggested approach in the numerical model is to either delay sampling or to penalize design locations that are too close to the injection site. It should be noted that in this example, as well as all others, the additional constraint that sampling must occur in both zones was included. This is done because in general, one cannot estimate parameters in a zone if no observations are taken in that zone. In a practical case this can be considered as simply incorporating engineering judgement (or prior information) into the design process.

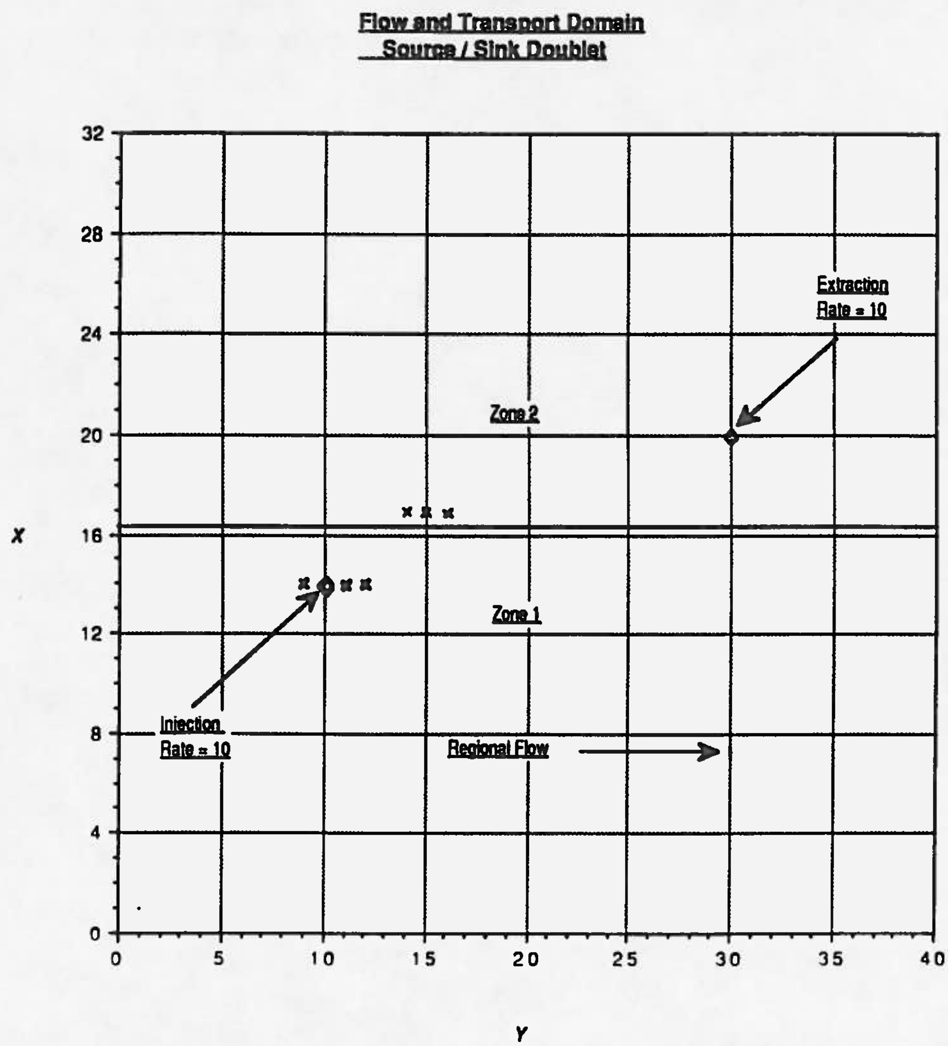

Figure 10 Optimal Design Case 1

10.2 Case 2

Case 2 is a homogeneous case, identical to Case 1 with a retardation factor of 2.0. Table 2 shows the parameters used

Table 2
Parameters for Case 2

	Zone 1	Zone 2	Weight
K_x	3.0	3.0	1.0
K_y	3.0	3.0	1.0
S	1.0	0.1	1.0
α_L	0.5	0.5	1.0
α_T	0.5	0.5	1.0
R	2.0	2.0	1.0

Figure 11 shows the CR-objective surface for Case 2. Again there is a maximum at the source, however, the objective surface is different than in Case 1 with the high information zone less spread vertically. An obvious interpretation is that increased retardation delays lateral spreading until beyond the duration of the experiment.

Figure 12 shows the optimal six point design for this case. The design is somewhat different, with the zone 1 points moved forward and the zone 2 points moved back. The design although somewhat uninteresting makes sense. These two cases illustrate that the CR-objective criterion and the 0-1 knapsack algorithm can identify reasonable designs.

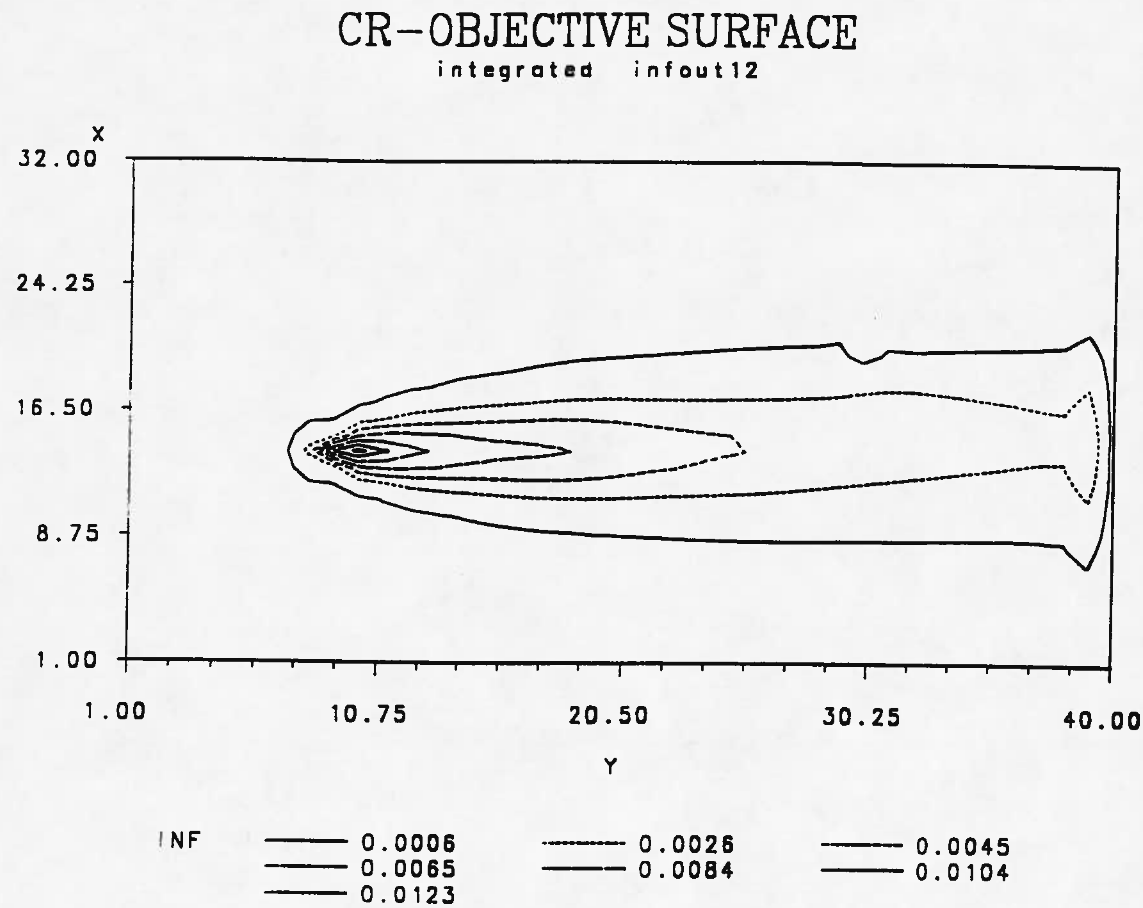

Figure 11 Objective Surface Case 2

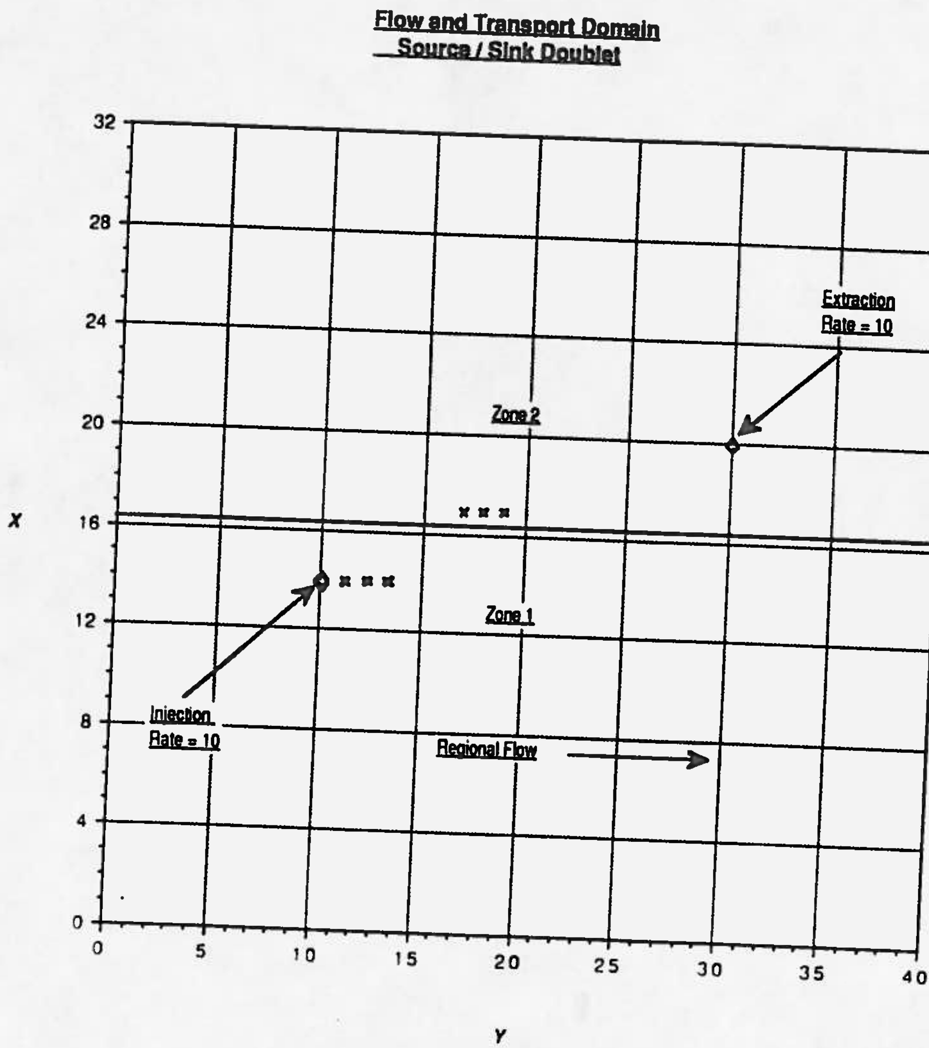

Figure 12 Optimal Design Case 2

10.3 Case 3

Case 3 explores a piecewise inhomogeneous case with the same doublet as before but with the parameters shown in Table 3.

Table 3
Parameters for Case 3

	Zone 1	Zone 2	Weight
K_x	1.5	3.0	1.0
K_y	1.5	3.0	1.0
S	0.1	0.1	1.0
α_L	0.5	0.5	1.0
α_T	0.5	0.5	1.0
R	1.0	1.0	1.0

Figure 13 shows the CR-objective surface for this case. In this case the objective surface looks about like that of Case 1. However, the contour plot shows a second peak near the extraction well which explains the dip in the outer contour in the first two cases. The effect of inhomogeneity in conductivity is negligible in this case, although it will be apparent in the next case. The effect is to turn the contours toward the extraction well, as well as shorten their extent in the slow zone.

Figure 14 shows the optimal six point design. The design is practically that of the previous (retarded) case for zone 1 while for zone 2 it is the average of the retarded case (Case 2) and the original case (Case 1).

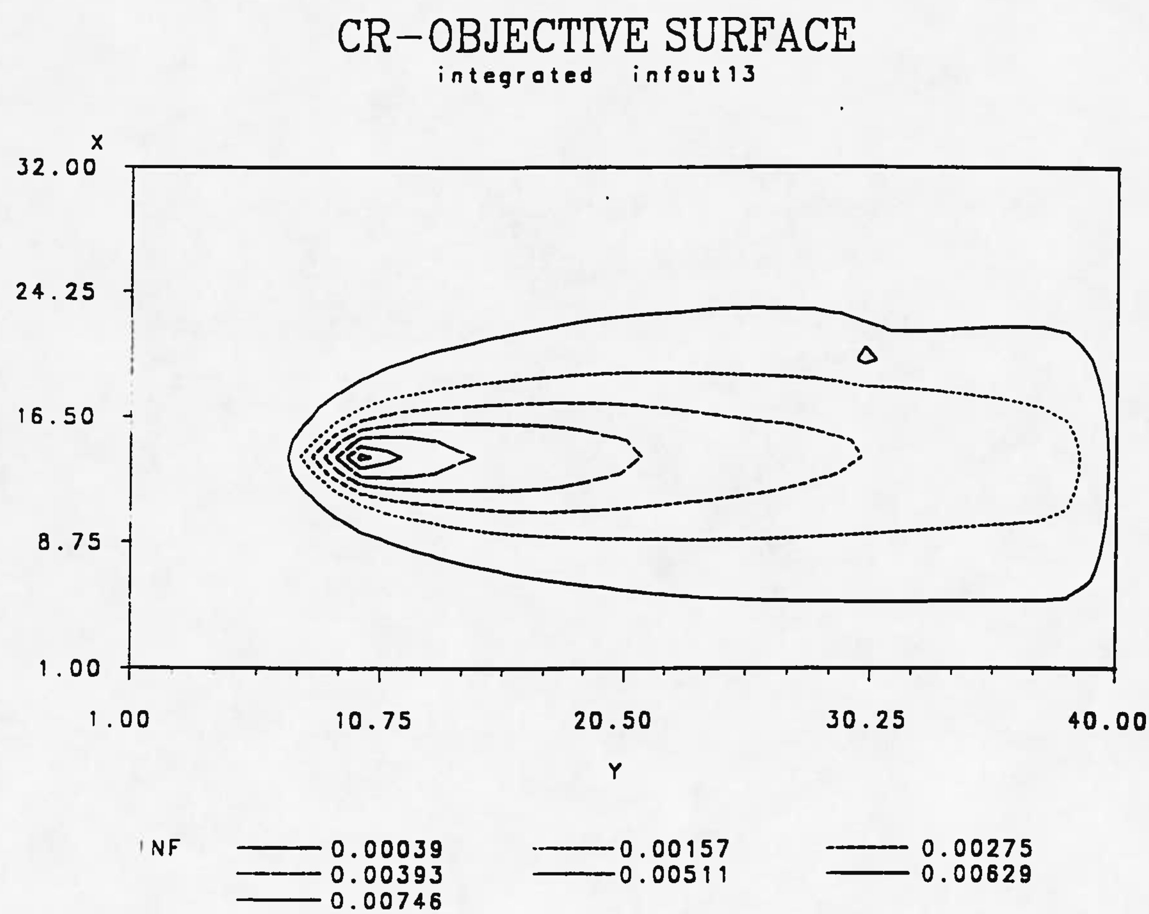

Figure 13 Objective Surface Case 3

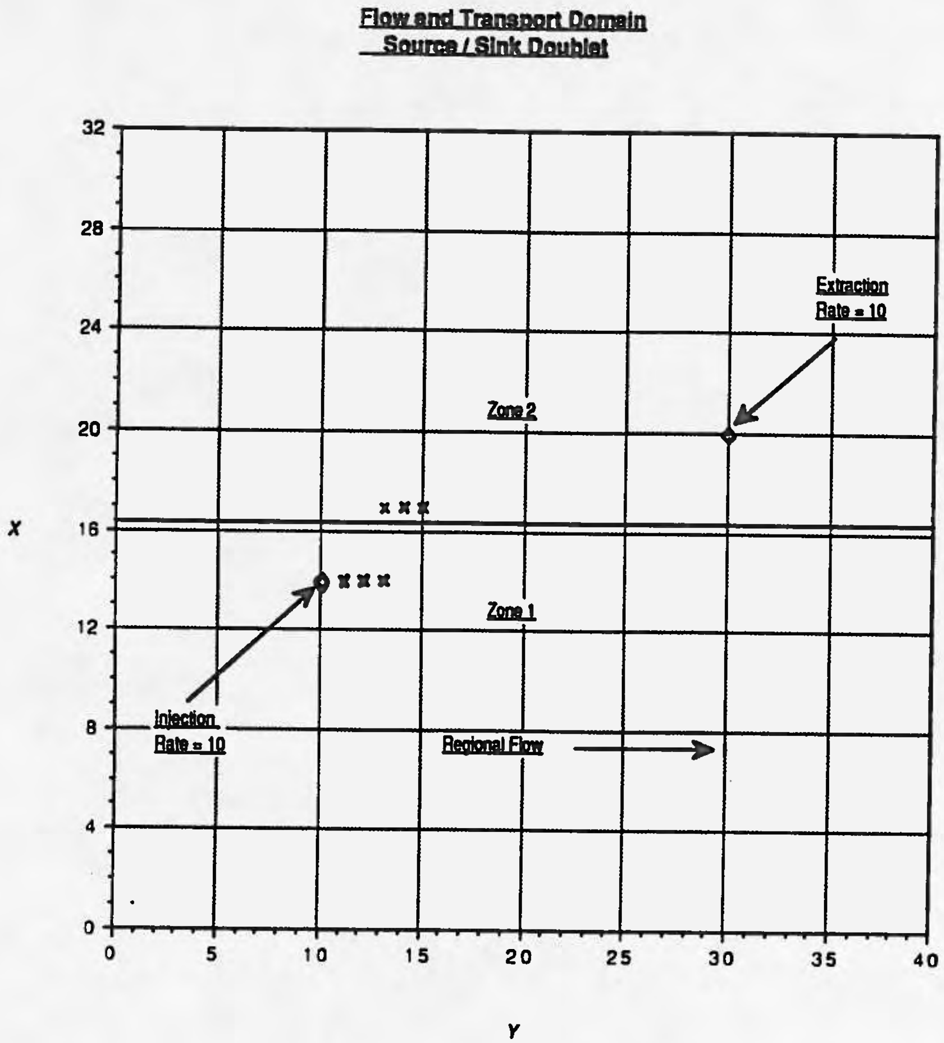

Figure 14 Optimal Design Case 3

10.4 Case 4

Case 4 is the same as Case 3 with retardation in both zones equal to 2.0 as in Table 4 below.

Table 4
Parameters for Case 4

	Zone 1	Zone 2	Weight
K_x	1.5	3.0	1.0
K_y	1.5	3.0	1.0
S	0.1	0.1	1.0
α_L	0.5	0.5	1.0
α_T	0.5	0.5	1.0
R	2.0	2.0	1.0

Figure 15 shows the CR-objective surface for this case. Like Case 2 it is nearly the same as the R=1.0 case except the contours are elongated. The effect of inhomogenity is more apparent, with the information contours more narrow in the fast zone.

Figure 16 shows optimal design.

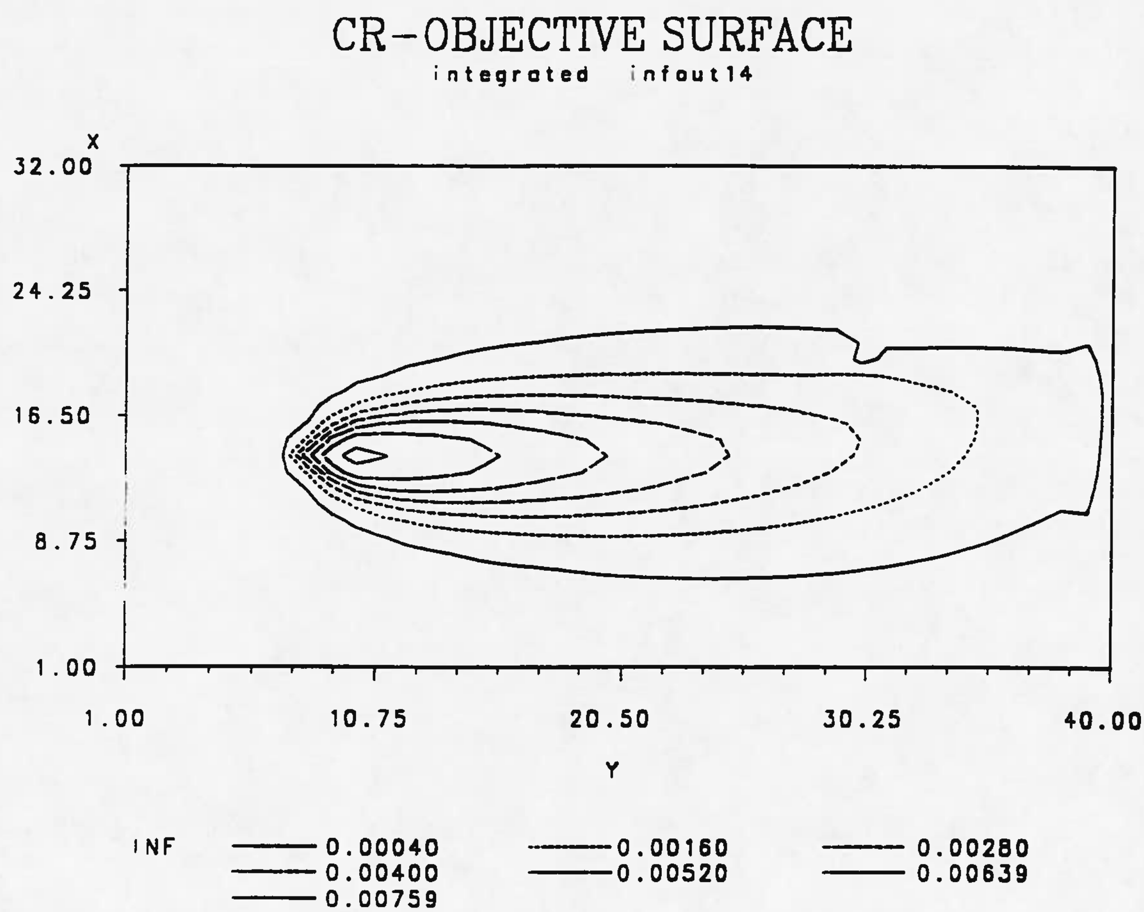

Figure 15 Objective Surface Case 4

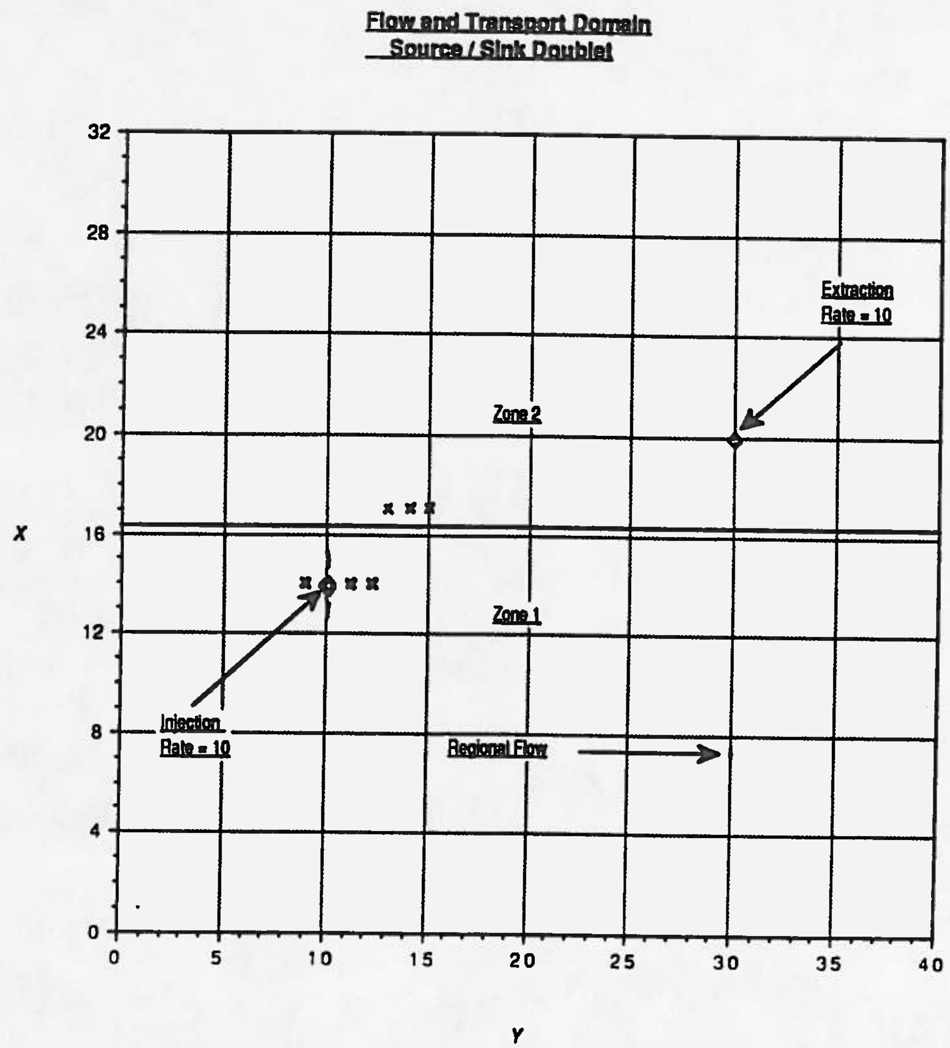

Figure 16 Optimal Design Case 4

10.5 Case 5

Case 5 and the next two cases explore a homogeneous case where the weights on the parameters are changed. Case 5 examines the case where only sensitivity and hence information on hydraulic conductivity is computed. Table 5 shows the parameters used.

Table 5
Parameters for Case 5

	Zone 1	Zone 2	Weight
K_x	3.0	3.0	1.0
K_y	3.0	3.0	1.0
S	0.1	0.1	0.0
α_L	0.5	0.5	0.0
α_T	0.5	0.5	0.0
R	1.0	1.0	0.0

Figure 17 shows the CR-objective surface. The maximal information occurs just downstream of the source. The explanation is that since information is mass dependent, points very near the injection site are most sensitive to changes in hydraulic conductivity. Since mass will still be concentrated at these points a small change in conductivity (and thus velocity) will cause a significant change in mass location over the experiment duration at these points. At other locations mass is already spread out so small changes have a negligible effect. Figure 18 shows the optimal six point design for this case. It is important to observe that the contour plot is nearly identical to that of Case 1.

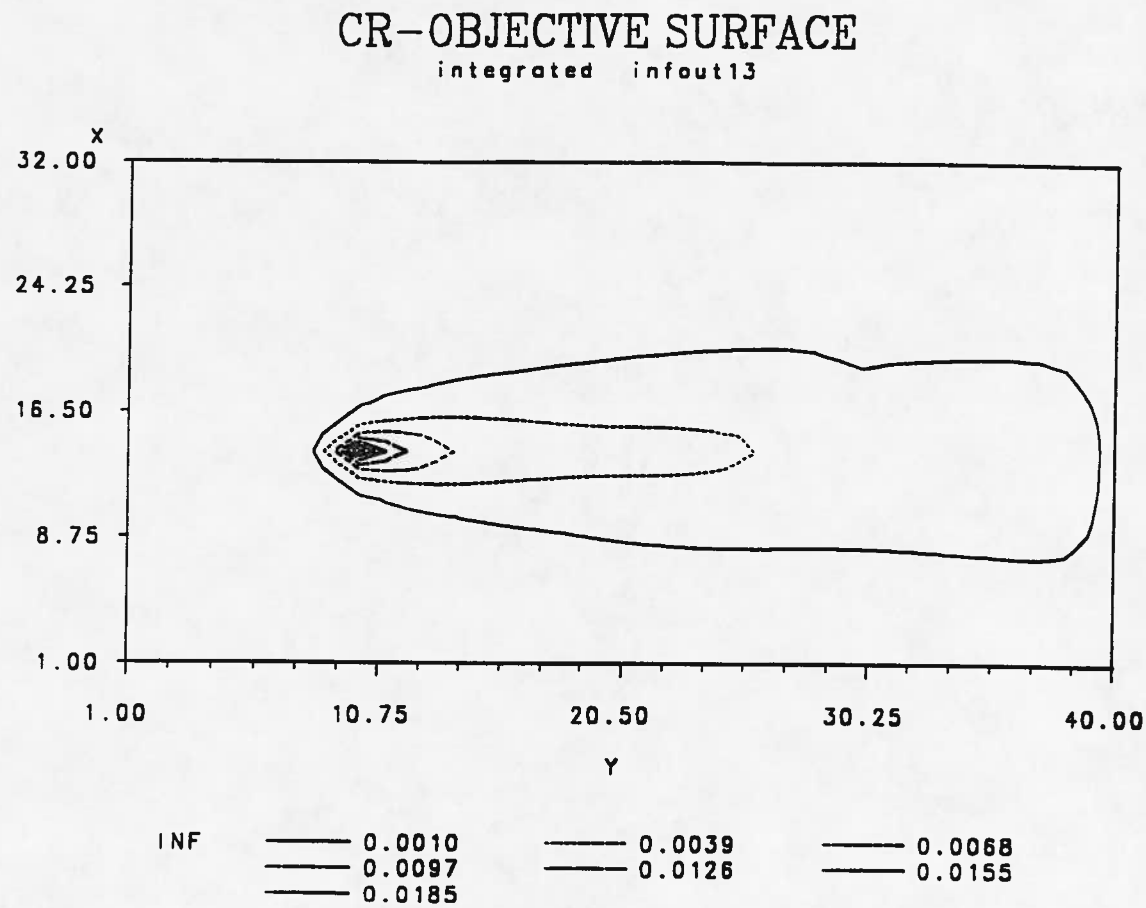

Figure 17 Objective Surface Case 5

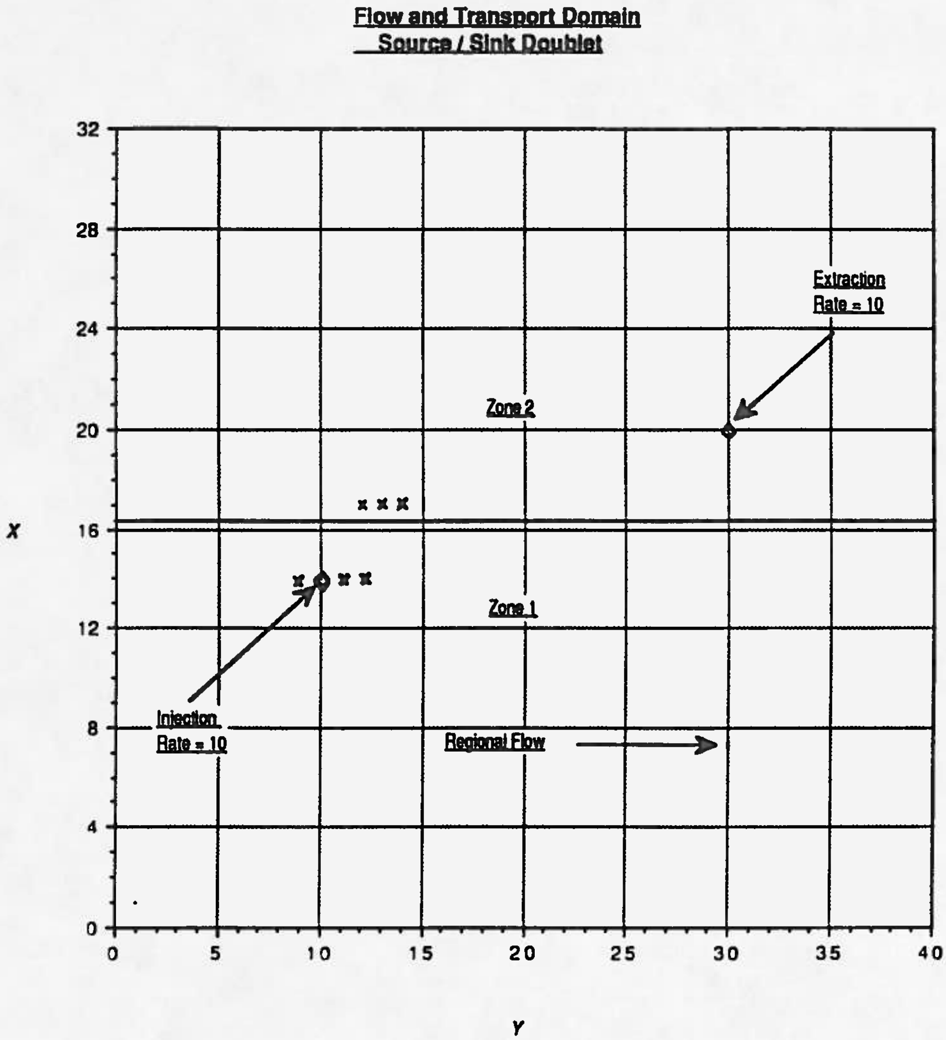

Figure 18 Optimal Design Case 5

10.6 Case 6

Case 6 is the same as before except the weights are zero everywhere except on α_L and α_T where they are set to one. Figure 19 shows the CR-objective surface. The plot is nearly the same as before, except less elongated in the regional flow direction. The contours are turned toward the extraction well, but it is barely noticeable. Again much of the information is contained near the source. The optimal design is shown in Figure 20.

The design is different than in Case 1 which indicates that in the doublet case, spreading terms are not dominant when all parameters are treated equally.

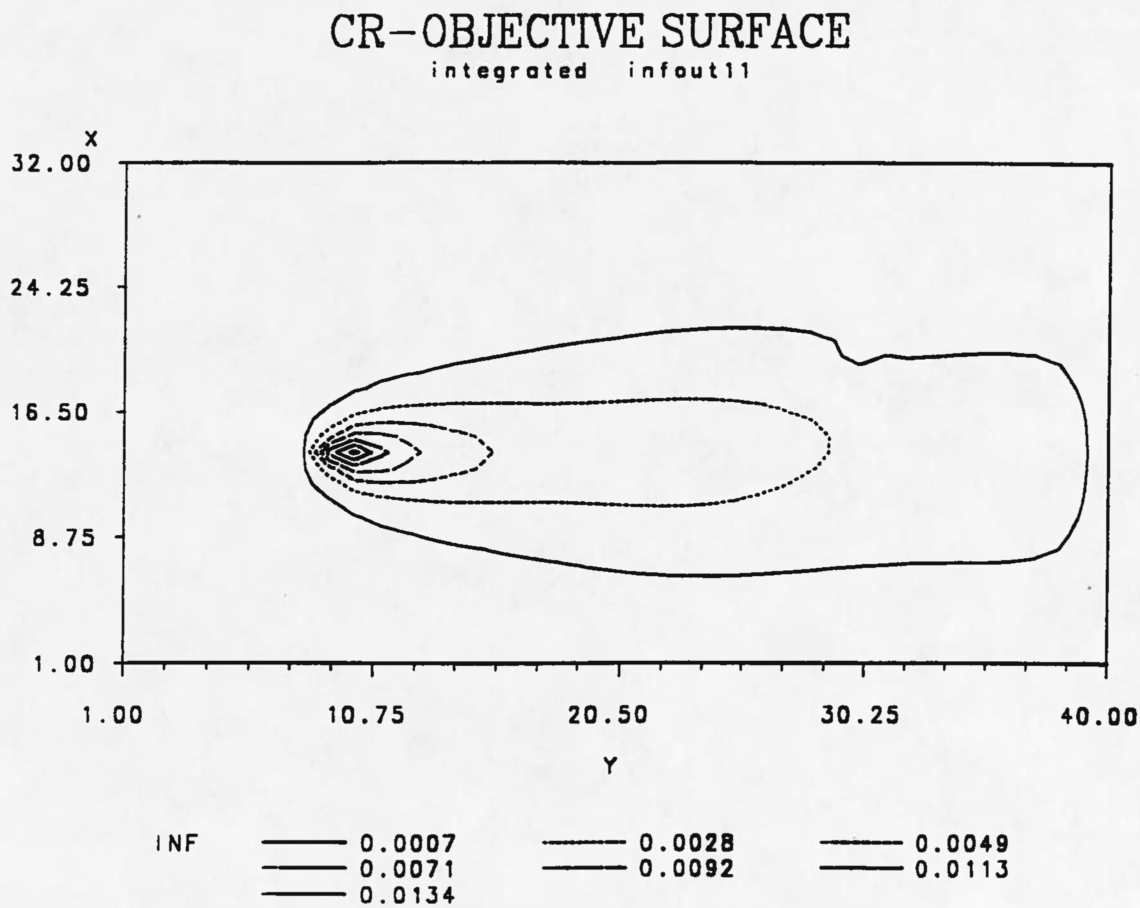

Figure 19 Objective Surface Case 6

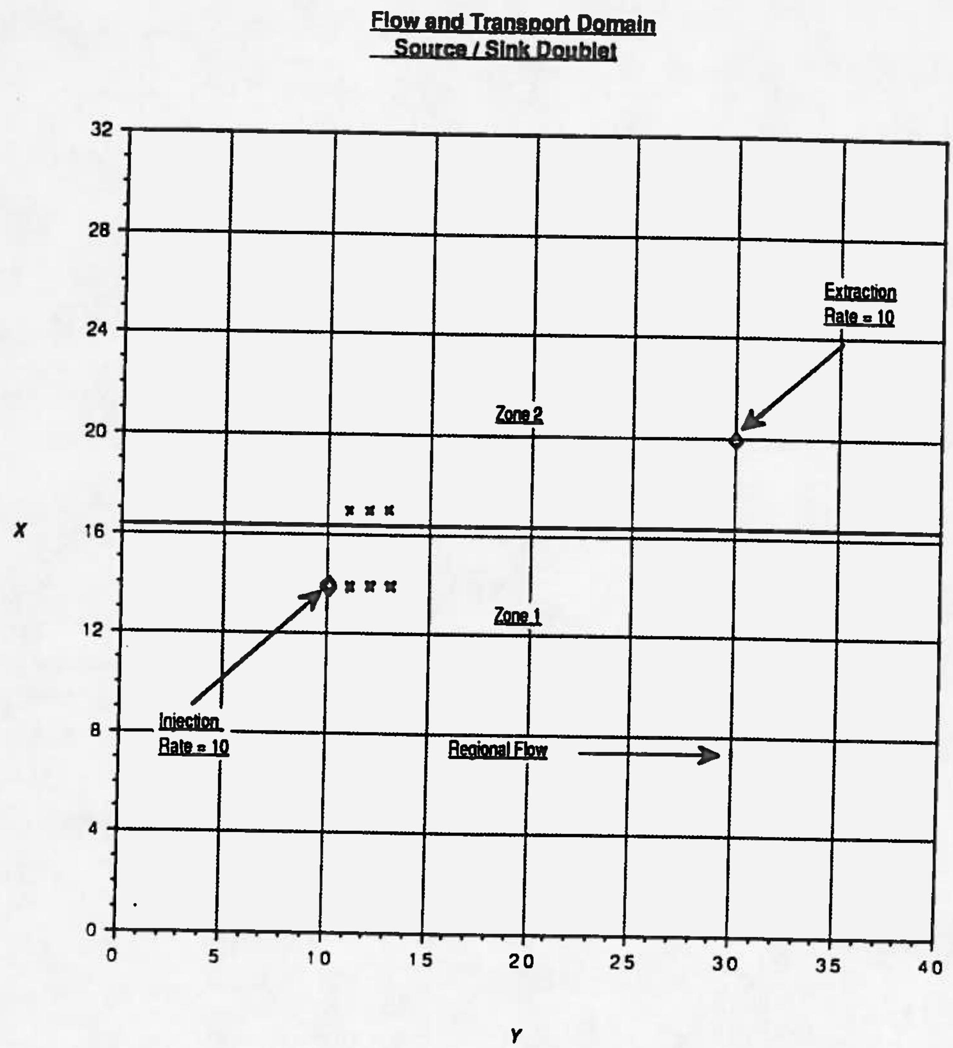

Figure 20 Optimal Design Case 6

10.7 Case 7

Case 7 is the same again, with only the non-zero weights of 1.0 placed on R. Figure 21 shows the objective surface contours which appear similar to the previous contour plots except the peak is more spread out. Figure 22 shows the optimal six point design.

The implications of these three cases on design is important. The transport parameters (dispersivity) have designs in one location, while the hydraulic parameters (conductivity) have designs at other locations. This illustrates the frustrating conclusion that designs that ensure good reliability about one parameter may be the worst possible for another.

Another point, in these experiments only concentration is measured, a head sensitive design might be able to handle the imbalance of location when one parameter is preferred to another. In Case 1 where all parameters are treated equally, the contour is the average in some sense of the contours of Cases 5 and 7. This makes sense since a concentration sensitive design is being used. Apparently the sensitivity to dispersivity does not dominate the solution, and these designs are determined by sensitivity to changes in transport parameters.

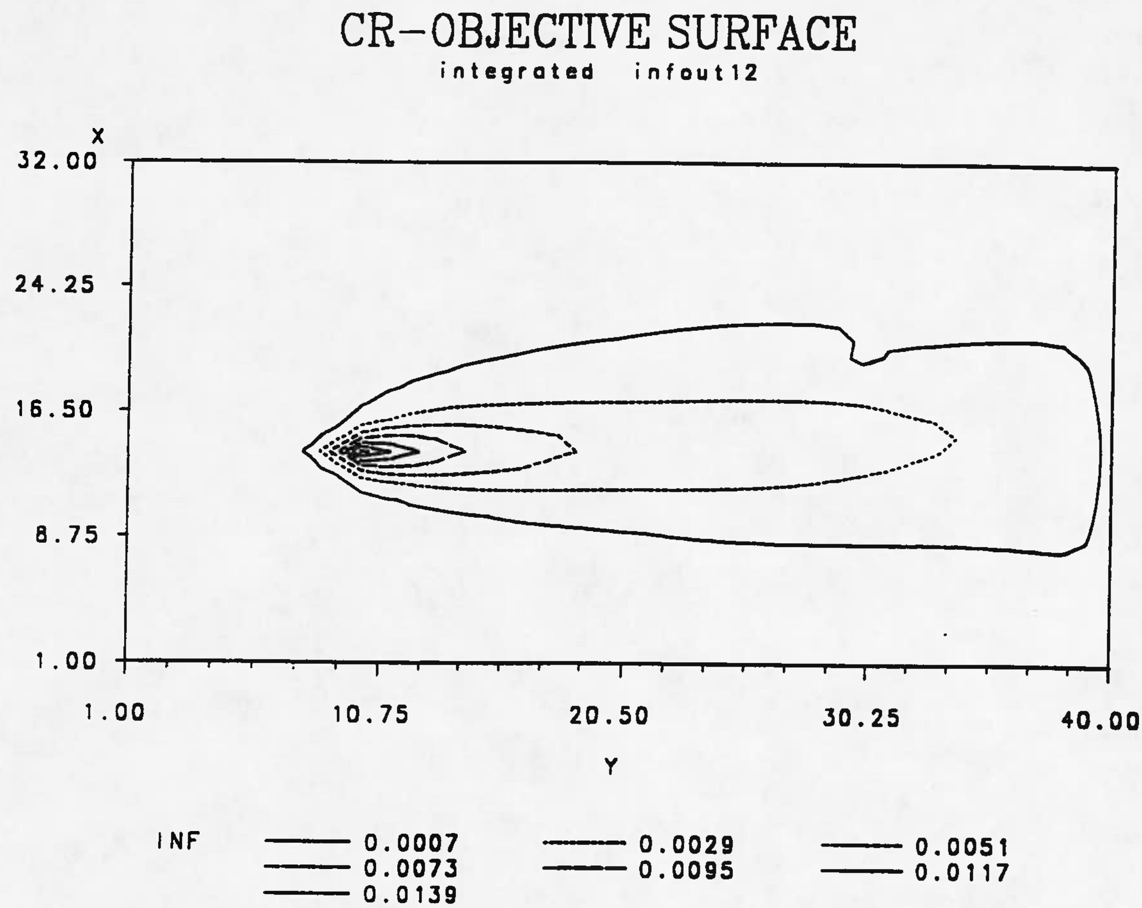

Figure 21 Objective Surface Case 7

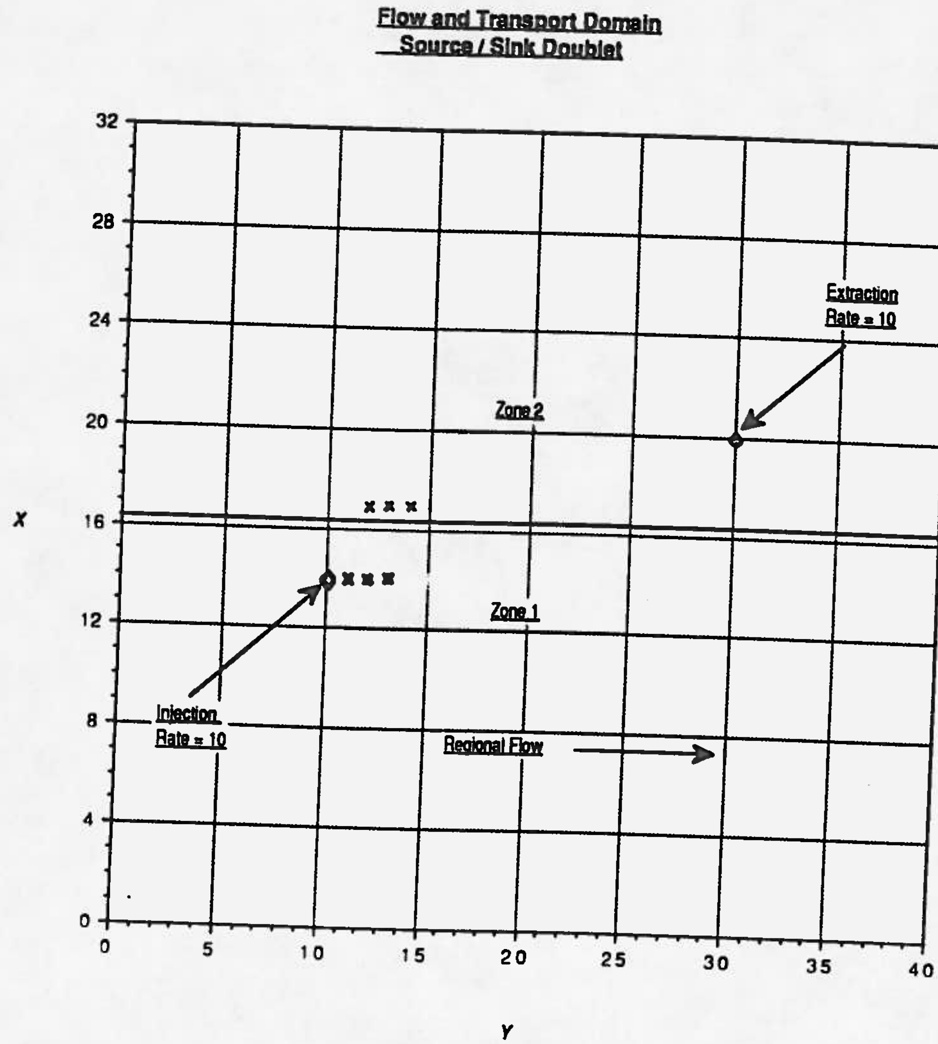

Figure 22 Optimal Design Case 7

10.8 Case 8

Case 8 examines a different flow domain. Injection is now along a line as is extraction. The injection rate is 36 m^3/day. Figure 23 shows the experiment configuration. This setup is to simulate a vertical flow problem where the placement of multilevel samplers is to be accomplished. The selection algorithm is modified to first select a line upon which to sample, then choose levels at which to sample along the line are given in Table 6.

Table 6
Parameters for Case 8

	Zone 1	Zone 2	Weight
K_x	3.0	3.0	1.0
K_y	3.0	3.0	1.0
S	0.1	0.1	1.0
α_L	0.5	0.5	1.0
α_T	0.5	0.5	1.0
R	1.0	1.0	1.0

Figure 24 shows the objective surface contours. The region of high information is located near the source as in previous examples. This is problematic in that the design is not too practical. The results make sense mathematically which suggests that the conceptual model for optimal design needs rethinking.

In practice a scaling effect has been observed that suggests one not sample near a source or too early in time. One approach that could be used would be to introduce a scale parameter that describes the averaging distance we would like our parameters to use, and then penalize any design points that are chosen too close to

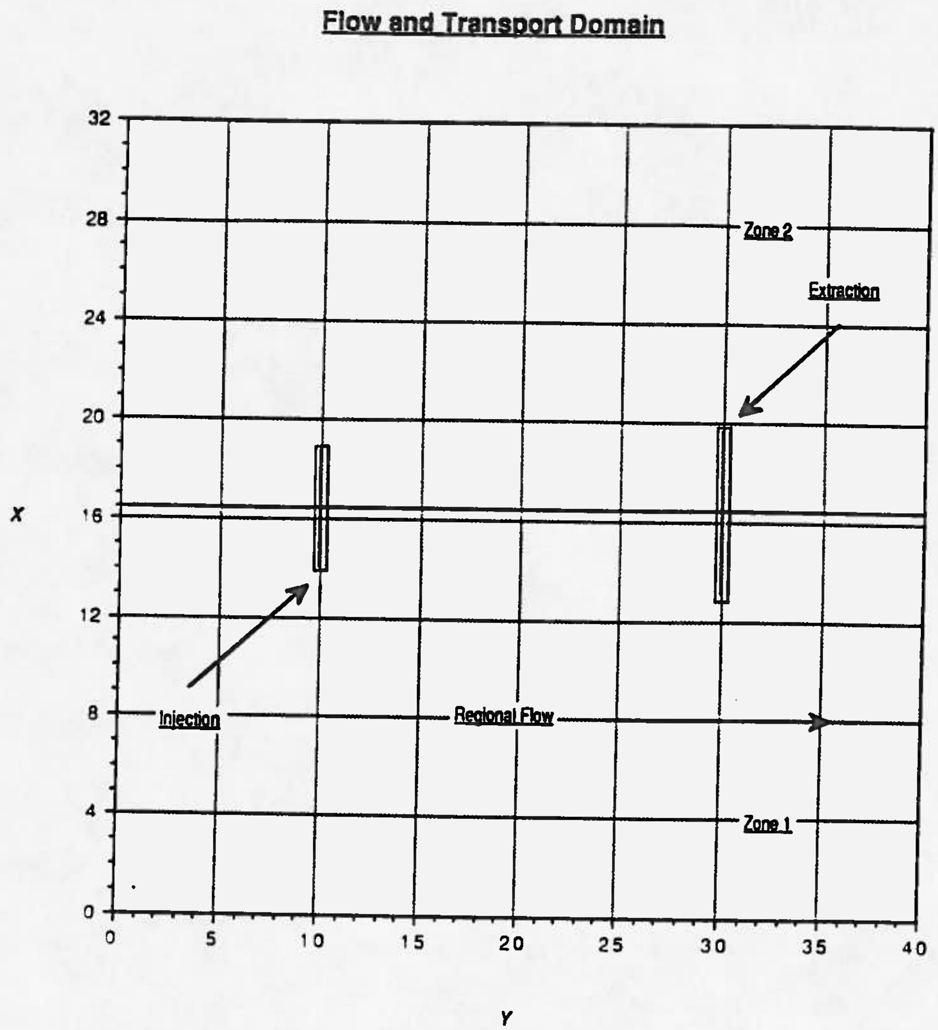

Figure 23 Flow and Transport Domain (Cases 8-15)

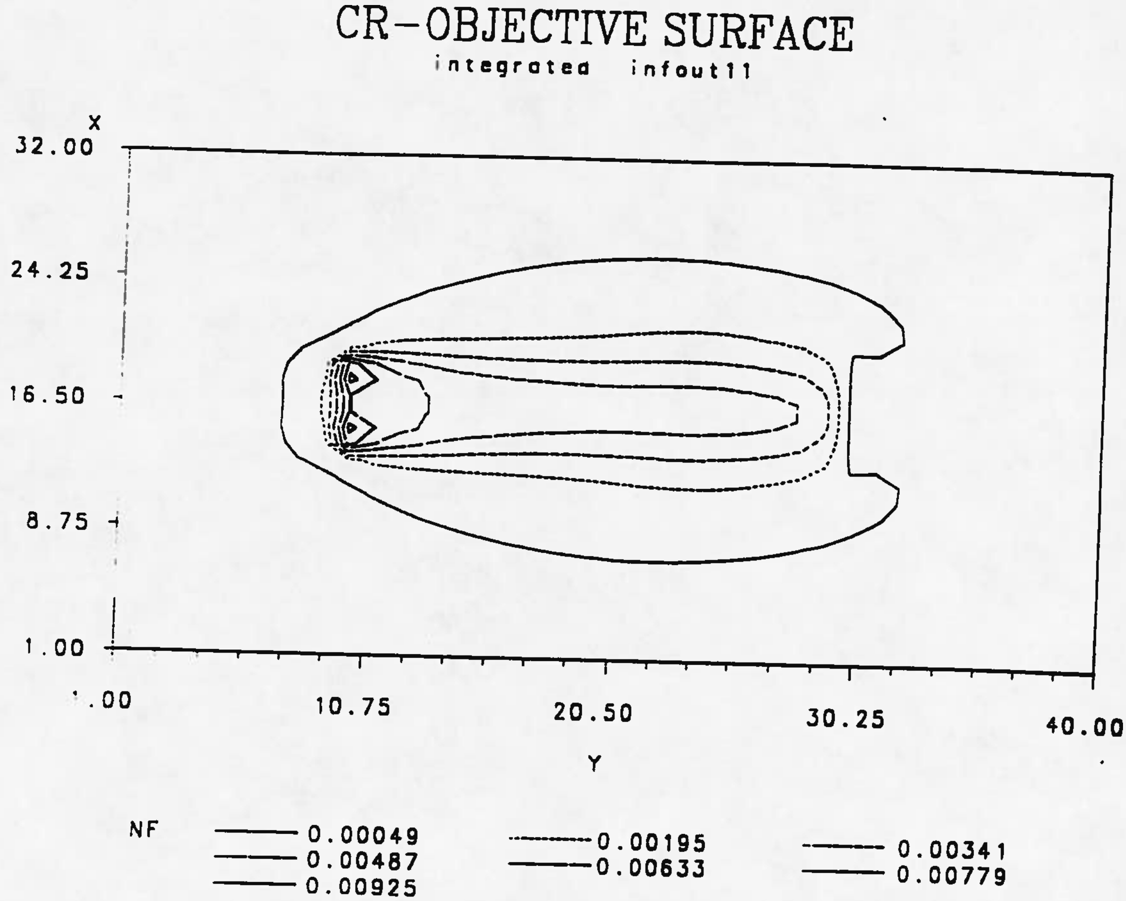

Figure 24 Objective Surface Case 8

the injections site.

Figure 25 shows the optimal multilevel six point design. The explanation for the two points that are displaced on a line toward the extraction site is that some estimate of curvature at each sample time must be available to estimate dispersivities.

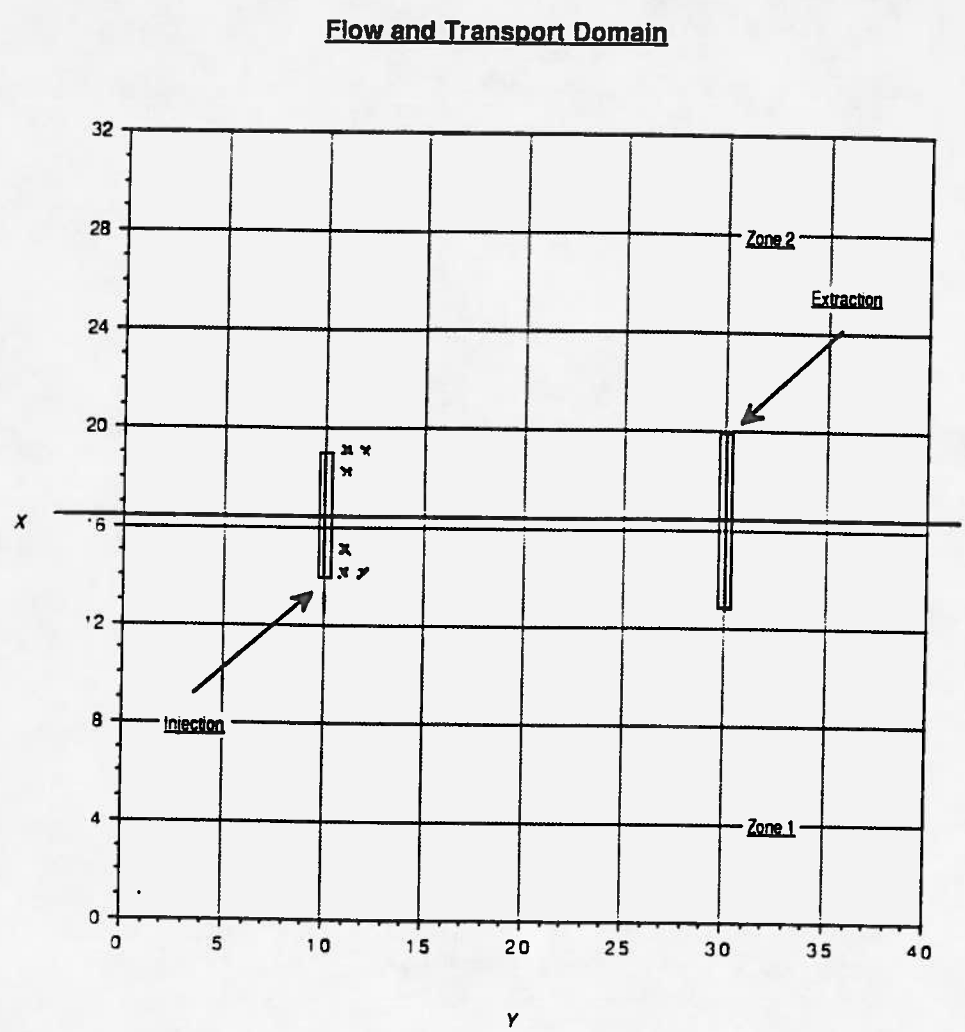

Figure 25 Optimal Design Case 8

10.9 Case 9

In Case 9 and the next two, the effect of changing the parameter weights is investigated. The parameters for Case 9 are the same as Case 8 except for the weights which are zero for all parameters but K_x and K_y.

Figure 26 shows the objective surface contours for this example. The region of high information looks the same as before (almost). It is observed that sensitivity to hydraulic conductivity only leads to a slightly different design than before. Figure 27 shows the design.

The difference is consistent with the nature of estimating conductivity, which would be done based on arrival and departure times of the pulse of mass. Since curvature of the breakthrough curve will be unimportant there is no need for sample points displaced in general flow direction.

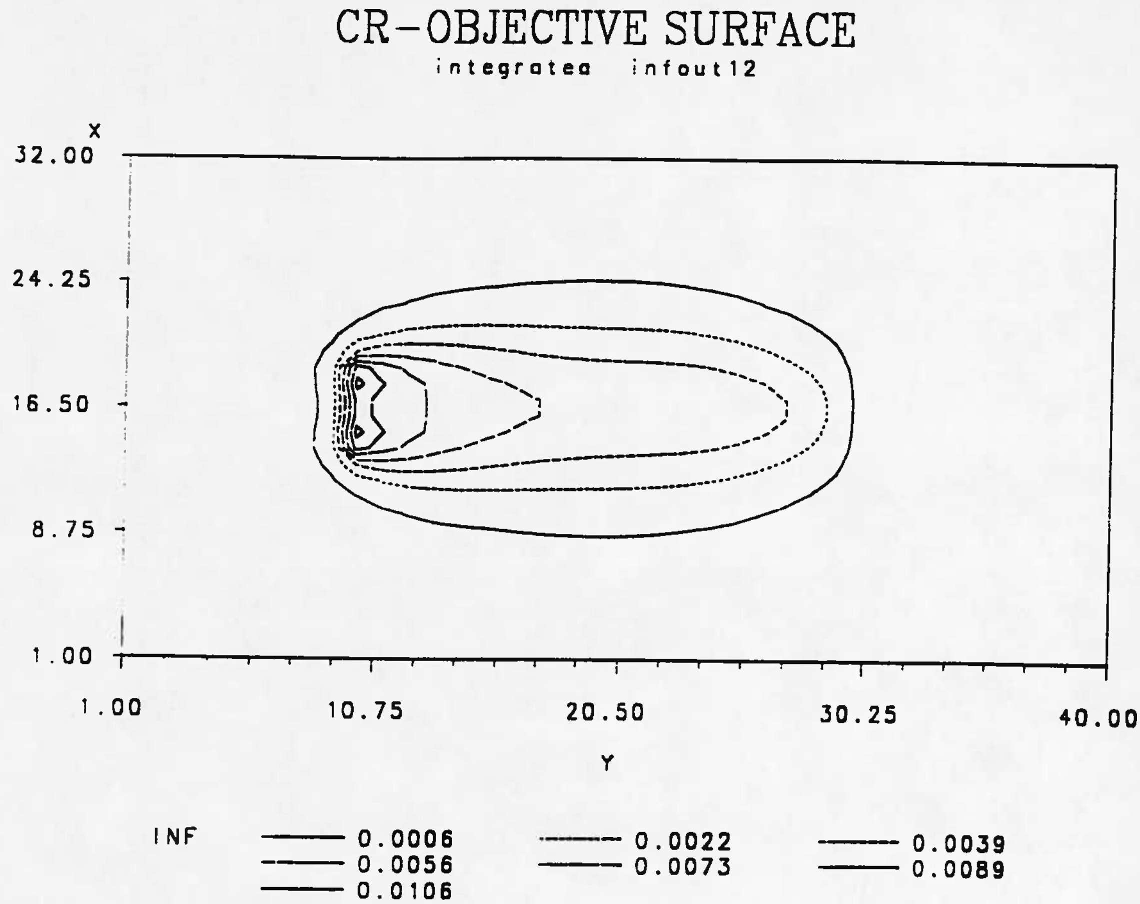

Figure 26 Objective Surface Case 9

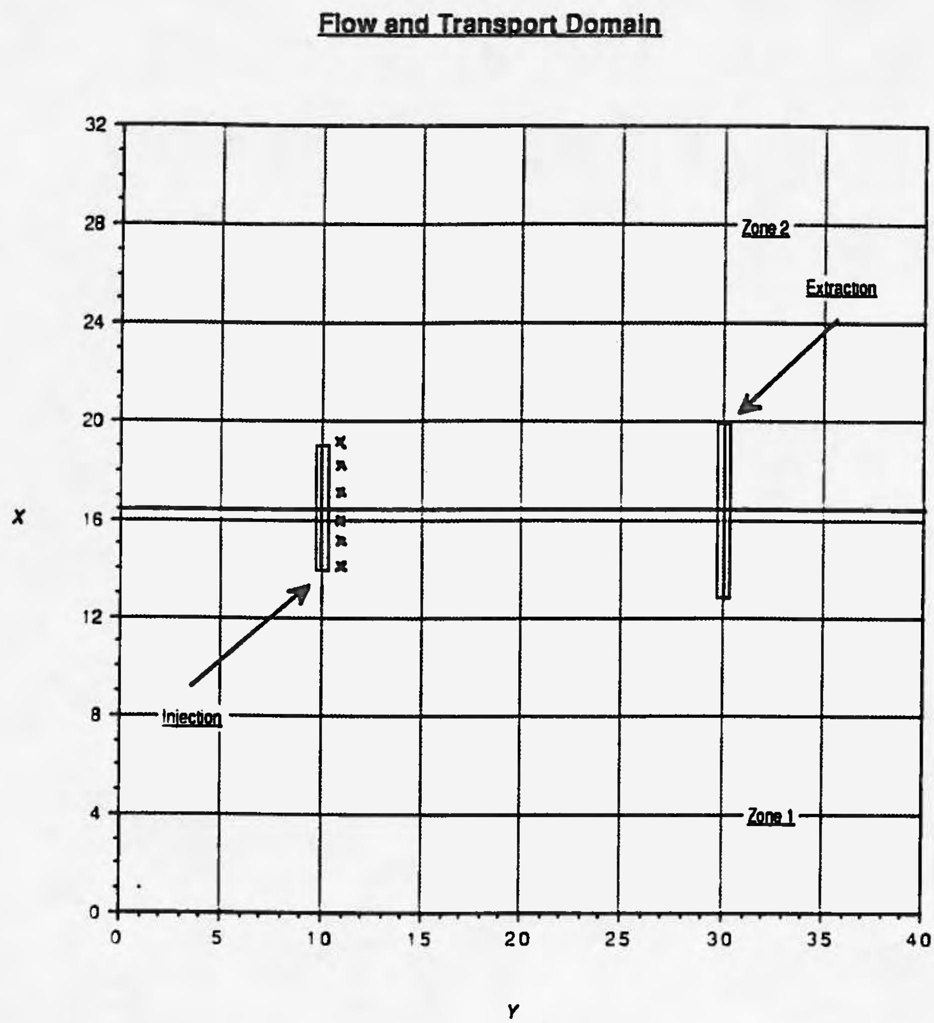

Figure 27 Optimal Design Case 9

10.10 Case 10

Case 10 changes the weights so that the non-zero weights are on α_L and α_T only. Figure 28 shows the objective surface contours.

The contour surface for this example is interesting. There are three regions of high information. One is just upstream of the injection site, one downstream a few nodes away, and the last near the extraction site. It is interesting to observe that all peaks are in line with the dominant streamline from source to sink.

The difference in appearance of this contour surface versus that of either Case 8 or Case 9 seems to indicate sensitivity to velocity is of different magnitude than that of dispersion.

In this example, when compared to the previous two cases, it is seen that velocity terms (i.e. hydraulic conductivity) completely dominate the objective surface. Clearly the choice of weights in the objective in these examples is poor from the standpoint of learning much about spreading.

Figure 29 shows the optimal multilevel design. Here the design is somewhat the reflection of the design for Case 9. The implication of this observation is that designs that are optimal for one parameter group will not be optimal for another. The other conclusion, at least for a practical case is that it will be necessary to design suboptimal with respect to hydraulics in order to learn about spreading.

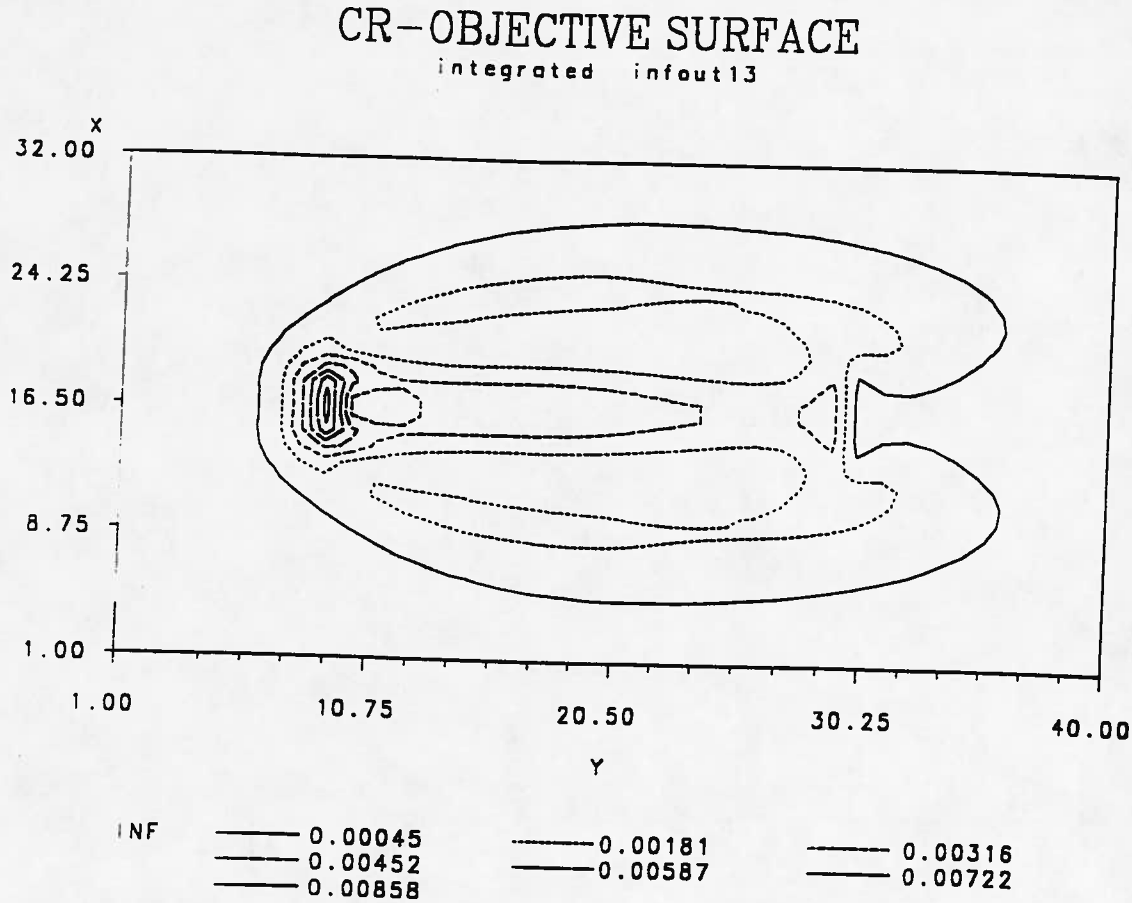

Figure 28 Objective Surface Case 10

Flow and Transport Domain

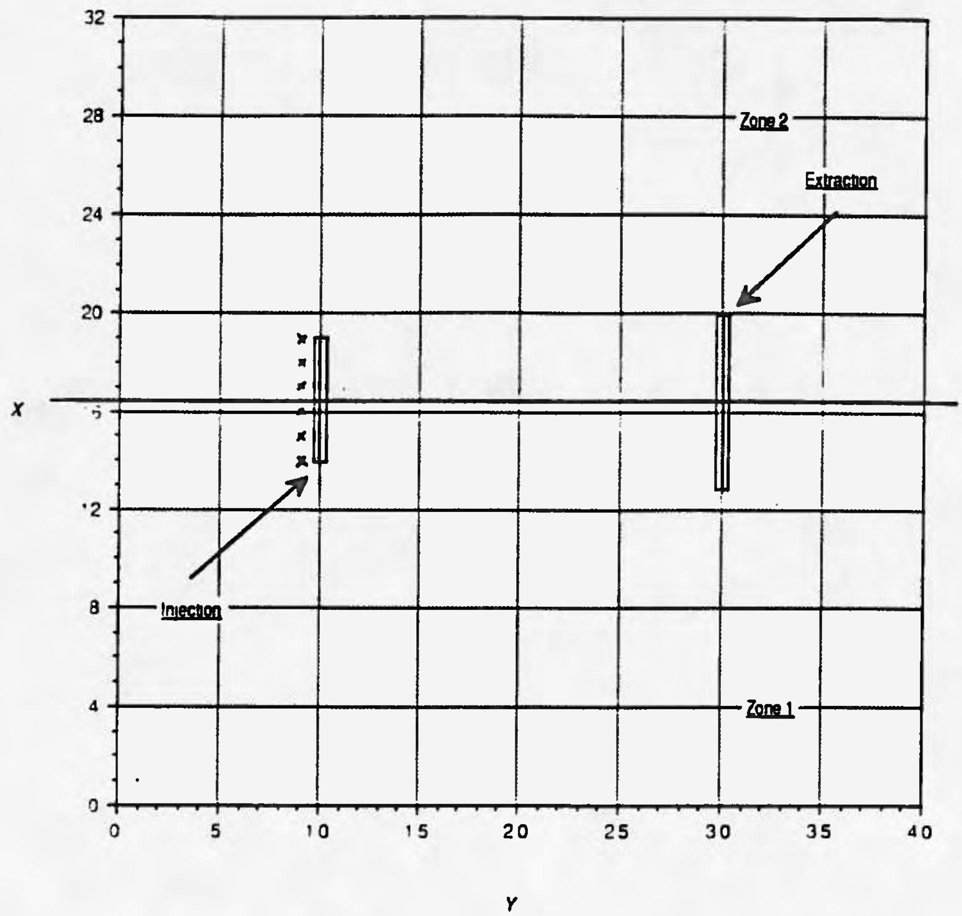

Figure 29 Optimal Design Case 10

10.11 Case 11

Case 11 changes the weights to non-zero on retardation only. The appearance is similar to Case 8 but the peak is more spread out. Figure 30 shows the objective surface contours.

Figure 31 shows the optimal multilevel design. In this case the design is identical to Case 8. This is interesting because it is observed that sensitivity to retardation is dominant in all these cases.

The implications of the general behavior are that the transport parameters and hydraulic parameters treated separately lead to different designs. Joint design depends on the importance of a particular parameter, since the uniform weight design usually mimics the retardation only design.

Knoppman and Voss (1987) found that the sensitivity to velocity is one order of magnitude larger than that to dispersion. This suggests adjusting the objective weights so that the various sensitivities (information) have about the same absolute magnitude at their respective peaks. Case 12 investigates this approach with the line source/sink design.

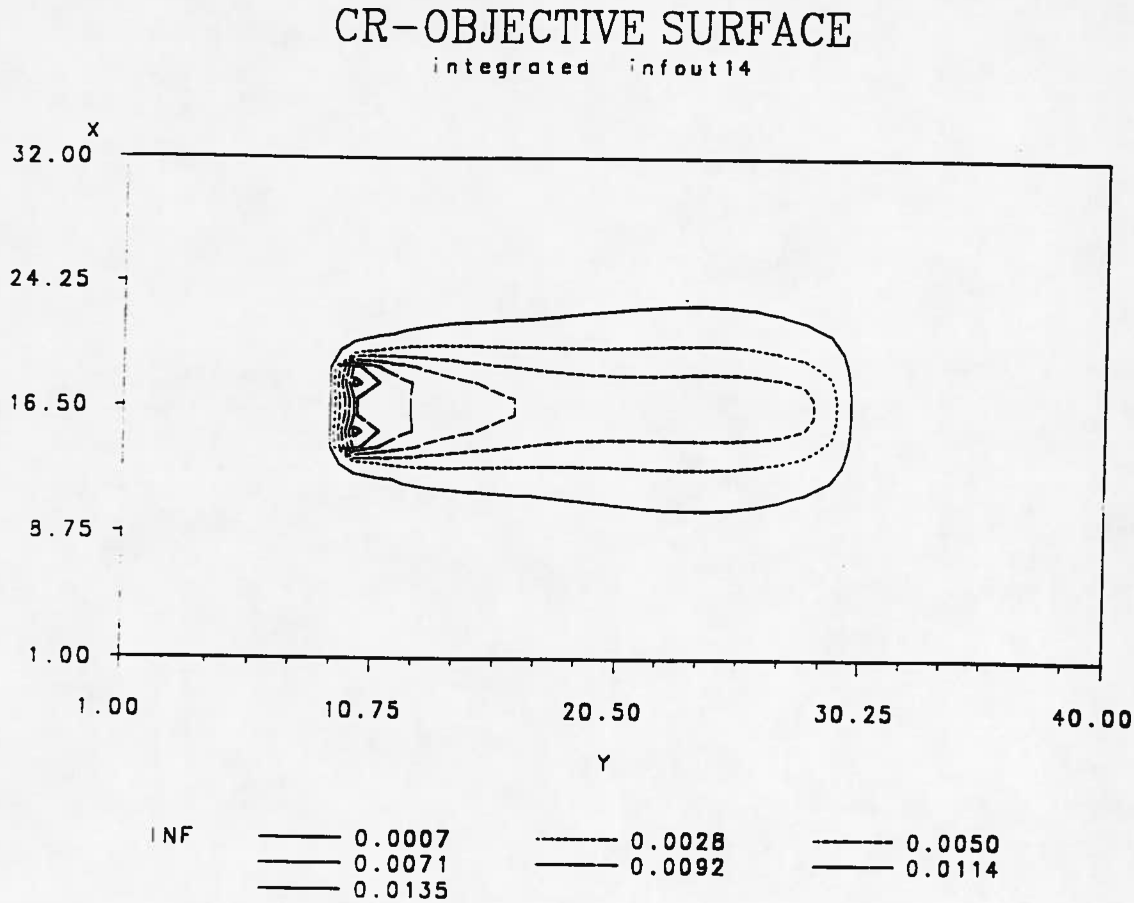

Figure 30 Objective Surface Case 11

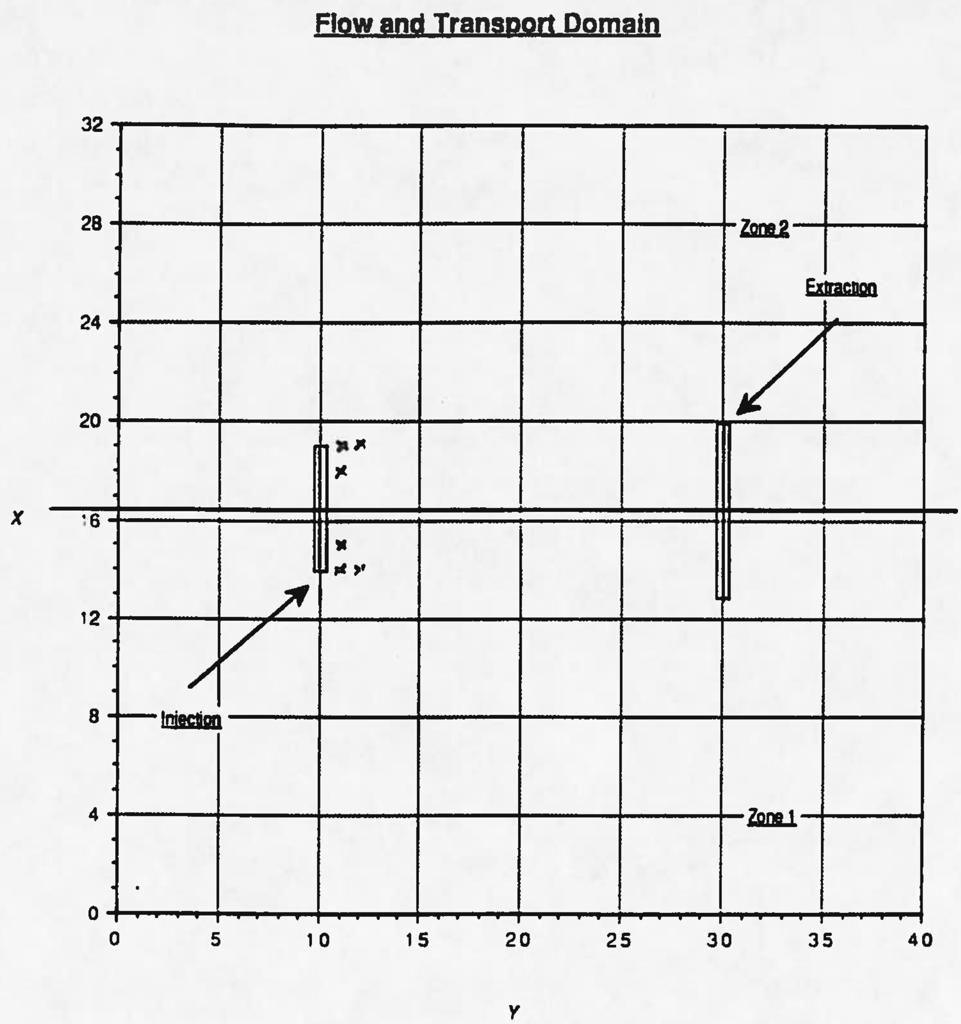

Figure 31 Optimal Design Case 11

10.12 Case 12

This case considers non-uniform weights in the objective function. The selection of non-uniform weights is necessary because the selection criterion gains computational simplicity at the expense of scale invariance (with respect to the relative sizes of uncertainity). The information available in Case 8 using the first 32 optimal points was 5.48. Case 9 where only conductivity is estimated the information is 0.44. In Case 10 where only diffusivity (dispersion) is considered the information is 0.83. Finally in Case 11 where retardation only is considered the information is 4.21. The sum of the partial information equals the uniform weight case (as in expected.) Normalizing the weights to reflect that dispersion terms are five times less dominant than retardation and fifteen times less than retardation leads to the weights shown in Table 7 below.

Table 7
Parameters for Case 12

	Zone 1	Zone 2	Weight
K_x	3.0	3.0	1.62
K_y	3.0	3.0	1.62
S	0.1	0.1	1.62
α_T	0.5	0.5	0.52
α_L	0.5	0.5	0.52
R	1.0	1.0	0.09

Figure 32 shows the objective surface contours. The change in weights has led to a more reasonable design in that the high information zone is now between the injection and extraction site. While there is no way of knowing whether these weights are optimal, making the relative magnitude of the squared sensitivities equal

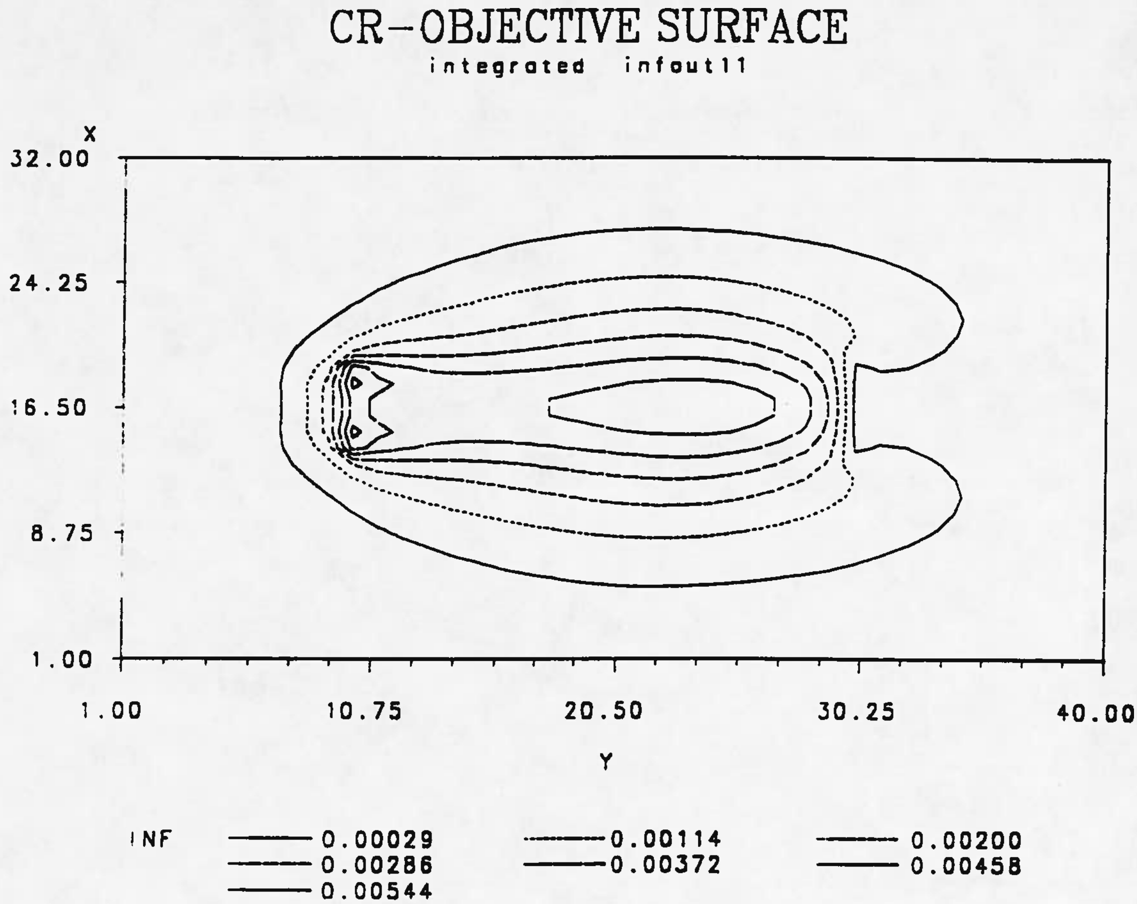

Figure 32 Objective Surface Case 12

leads to a more intuitively reasonable design. Clearly choice of weights does indeed change the design, but the actual weights used will depend on the experimental and confidence in the parameter estimates used to construct the design.

Figure 33 shows the optimal two location multilevel design for this case.

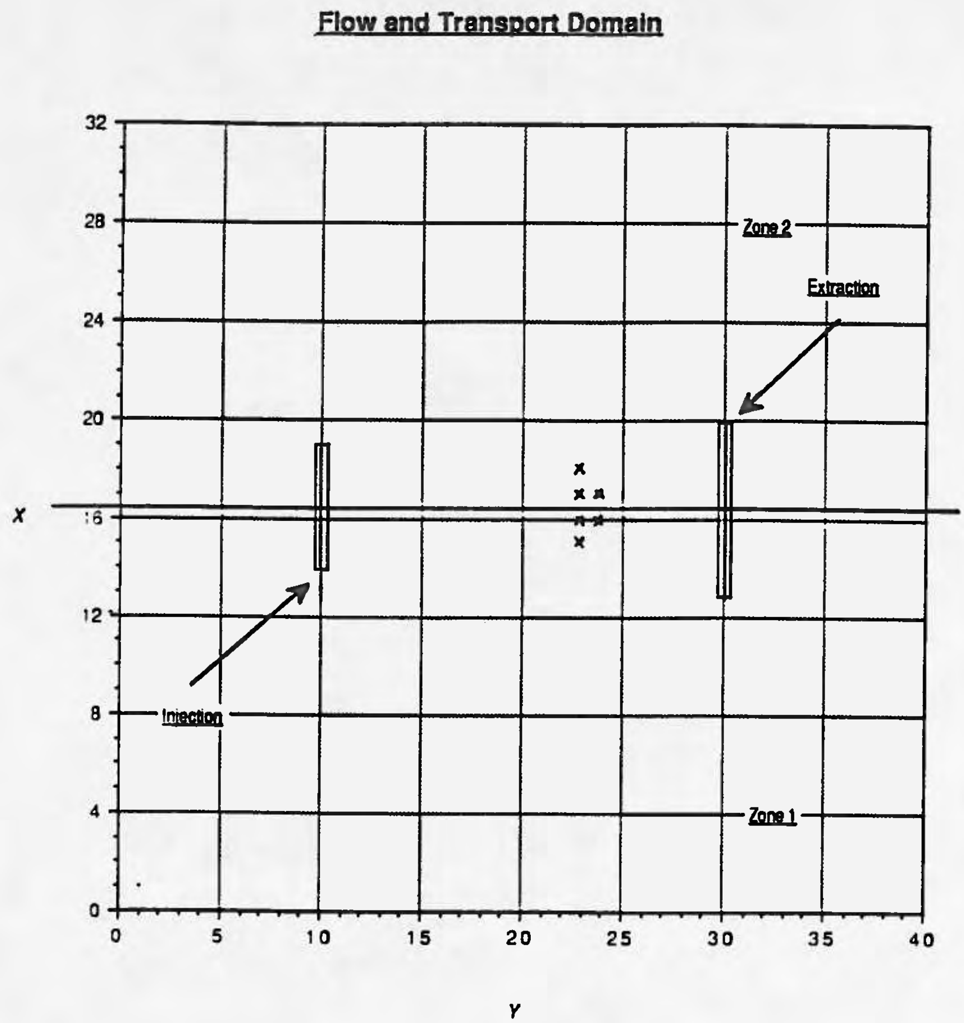

Figure 33 Optimal Design Case 12

10.13 Case 13

Case 13 explores the weighted objective with longitudinal dispersivity reduced. The parameters for Case 13 are shown in Table 8 below.

Table 8
Parameters for Case 13

	Zone 1	Zone 2	Weight
K_x	3.0	3.0	1.62
K_y	3.0	3.0	1.62
S	0.1	0.1	1.62
α_T	0.1	0.1	0.52
α_L	0.5	0.5	0.52
R	0.5	0.5	0.59

The objective contours are shown in Figure 34. In this case the maximum information locations are near the extraction site. The implications for sampling design are that an optimal design is sensitive to spreading parameters, and the most difficult to estimate parameters have the greatest effect on design. It is also important to observe that the peaks are off the dominant streamline indicating that when spreading transverse to the general flow is more significant than parallel to the flow, samples must be made transverse to the dominant streamline to obtain information about the parameters.

Figure 35 shows the optimal six point design for this case.

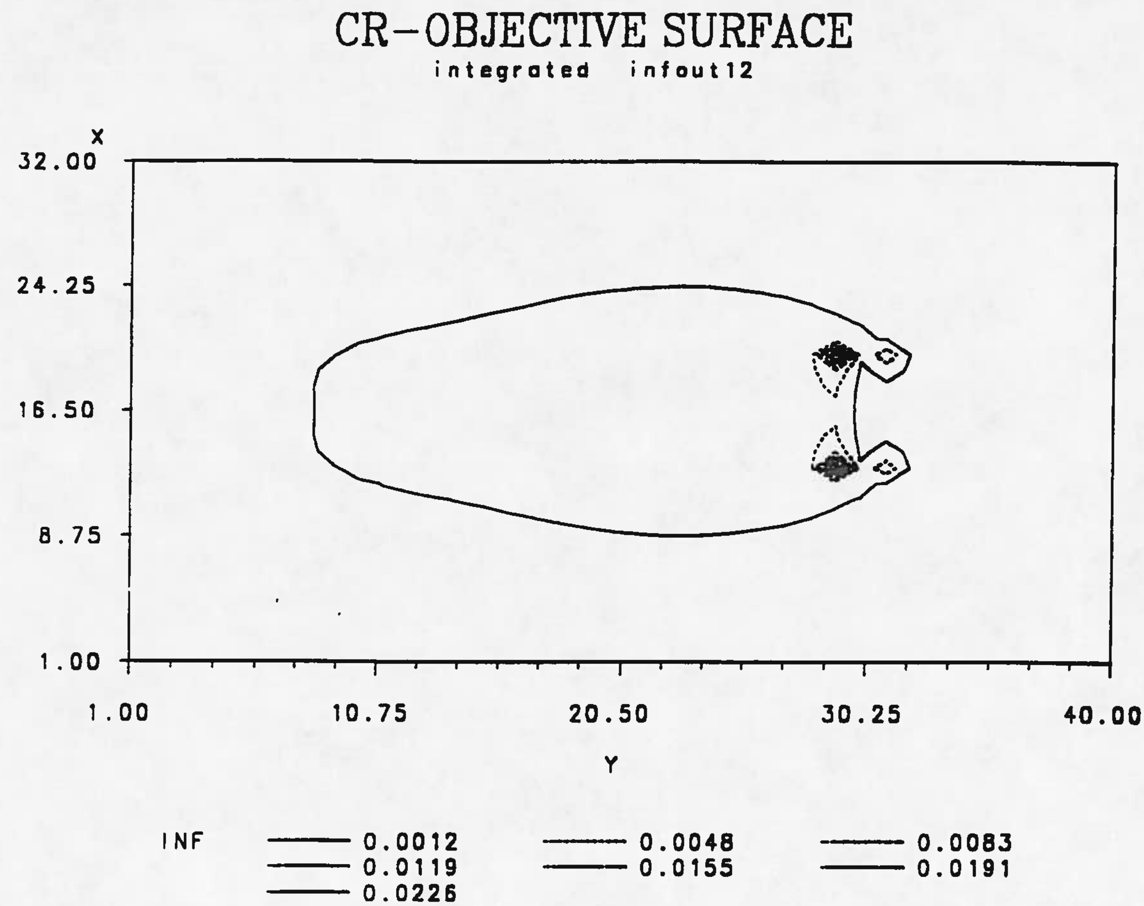

Figure 34 Objective Surface Case 13

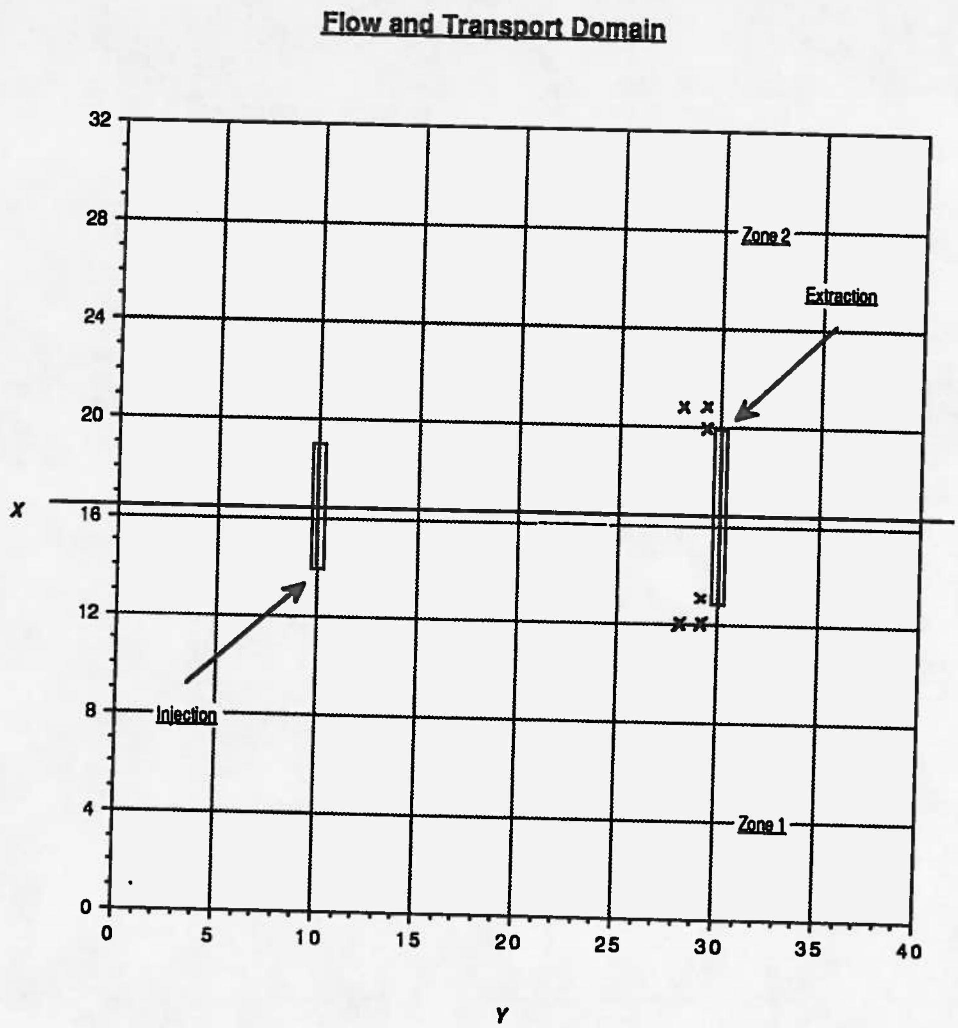

Figure 35 Optimal Design Case 13

10.14 Case 14

Case 14 is an inhomogeneous case with sampling delayed for five days. Figure 36 shows the contour surface for this case. The parameters used are shown in Table 9 below.

Table 9
Parameters for Case 14

	Zone 1	Zone 2	Weight
K_x	1.5	3.0	1.62
K_y	1.5	3.0	1.62
S	0.1	0.1	1.62
α_T	0.5	0.5	0.52
α_L	0.5	0.5	0.52
R	1.0	1.0	0.09

It is interesting to observe that the maximum information peak in the slow zone (zone 1). This makes some sense by extrapolating the argument for near source sampling in the uniform weight case, that is the mass is more concentrated in the slow zone.

The delay in sampling does not appear to affect the location of the maximum point with respect to the horizontal distance from the source. Figure 37 shows the optimal six point design.

All the examples presented so far have neglected one important fact, that is the parameters are unknown prior to conducting an experiment. It is a circular problem, probably best handled by a sequential design approach. Stated succinctly, the parameters are required to identify an optimal design in order to obtain the data to

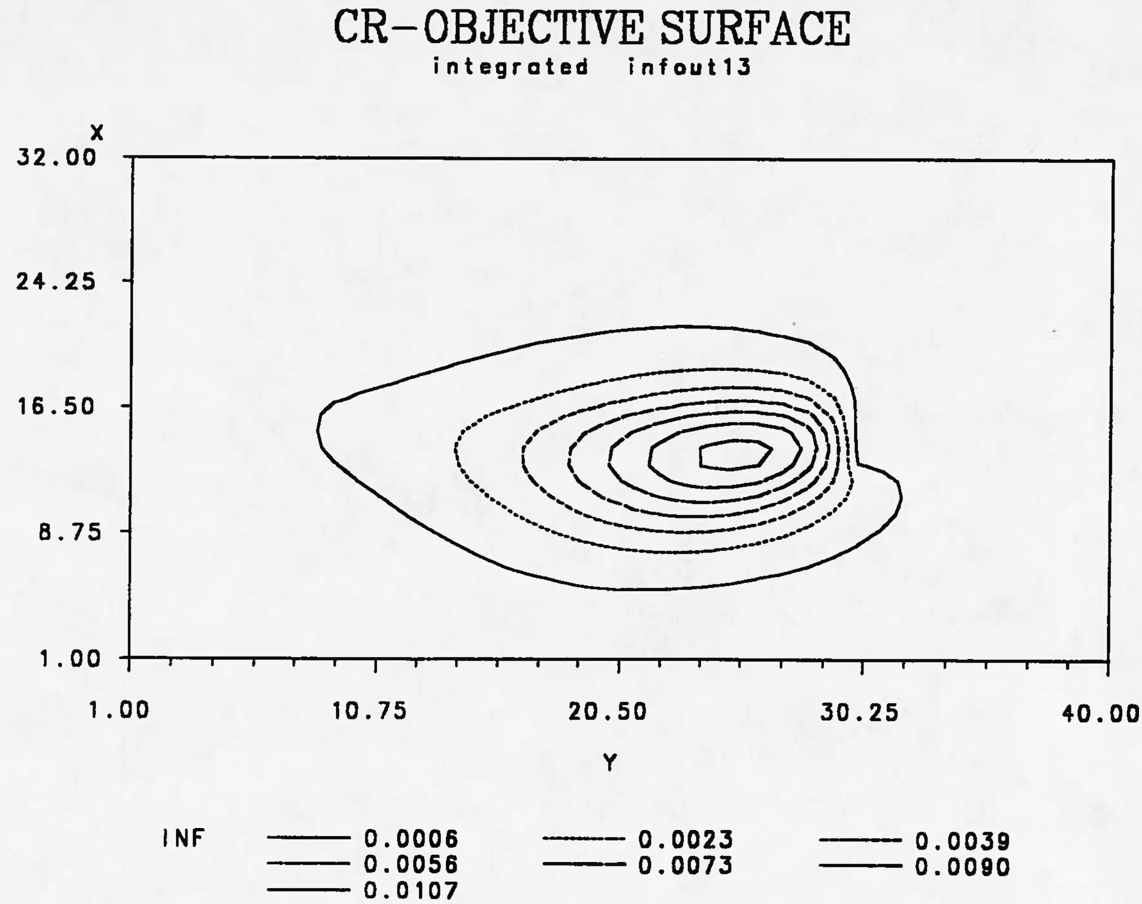

Figure 36 Objective Surface Case 14

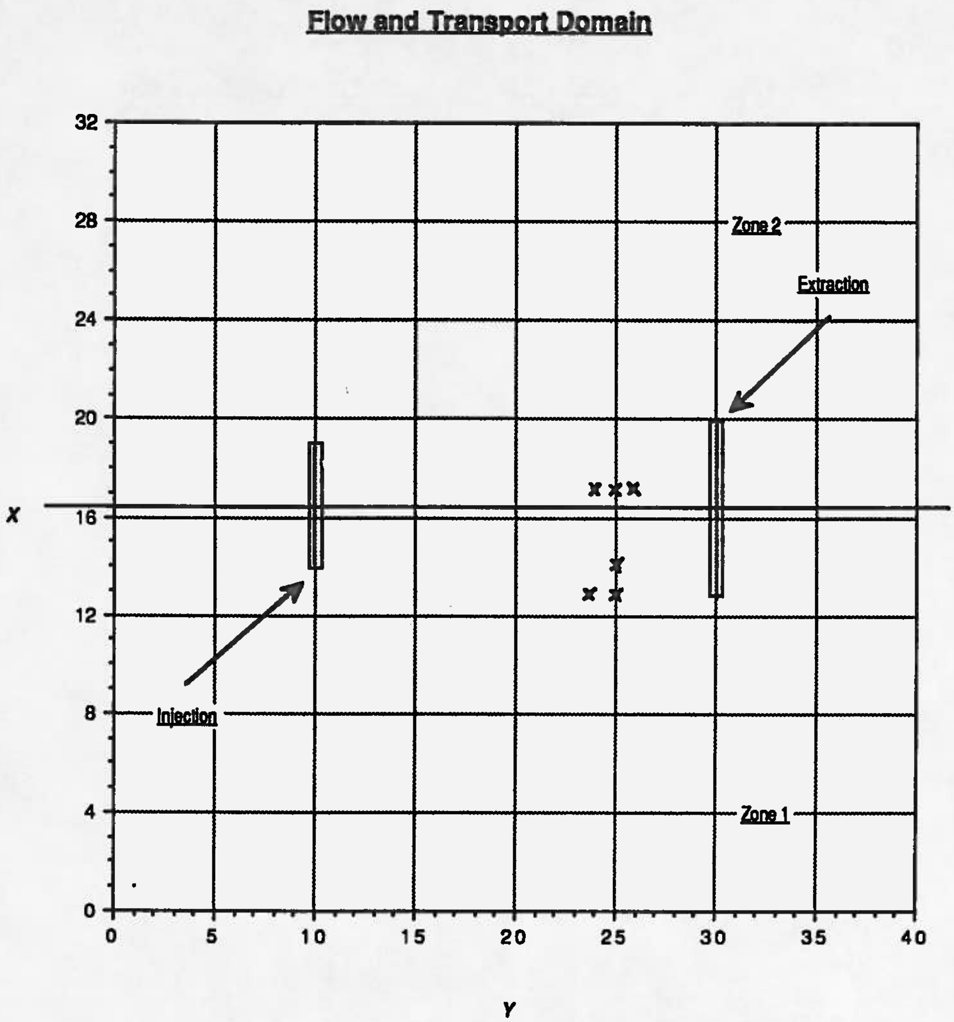

Figure 37 Optimal Design Case 14

estimate them, but they are unknown beforehand.

Two approaches that are used are to design an experiment that performs well in the average sense or to design an experiment that performs well in the min-max sense. D-optimality is a common criterion for these approaches although in general the approaches generate different designs (Walter and Pronzato, 1987). Although these approaches identify robust designs they are not independent of the underlying parameters (which are unknown) and thus are dependent on the estimates (which are used to approximate the covariance matrix). To gain insight in this respect, this study uses a Monte Carlo approach to identify parameter ranges for which a design remains optimal. Since it has been demonstrated that the parameters of interest change the objective surface in a complex fashion this trial-and-error approach should yield some insight to "robustize" a design.

The approach is the following, parameters are chosen from a uniform distribution with known upper and lower bounds. Several thousand realizations are performed for some budget level and the frequency of selected points is saved. With parameters at some setting and small ranges a unique design should appear with frequency 1.0. The ranges are adjusted until this frequency changes. This is then said to be the optimal design for all parameters in the stated ranges. An example is shown.

10.15 Case 15

This is an inhomogeneous case where the parameters are chosen from uniform distribution centered as in Table 9.

Table 10
Parameters Ranges for Case 15

	Zone 1	Zone 2
K_x	1.0 - 2.0	2.5 - 3.5
K_y	1.0 - 2.0	2.5 - 3.5
S	0.9 - 1.1	0.9 - 1.1
α_L	0.25 - 0.75	0.25 - 0.75
α_T	0.25 - 0.75	0.25 - 0.75
R	0.75 - 1.25	0.75 - 1.25

Table 10 shows the ranges for which the design based on Table 9 does not change.

Table 11
Parameters Ranges for Stable Design

	Zone 1	Zone 2
K_x	1.26 - 1.98	2.57 - 3.45
K_y	1.08 - 1.93	2.55 - 3.37
S	0.9 - 1.0	0.9 - 1.0
α_L	0.27 - 0.55	0.38 - 0.69
α_T	0.33 - 0.49	0.44 - 0.72
R	0.75 - 1.15	0.78 - 1.24

For parameters outside these ranges the marginal decrease in information by switching from the optimal design to an average design can be approximated by comparing the information in the average design with the information in the optimal design. On the average the information loss is dramatic, about 70%. This is alarming in that a one-shot approach will have serious difficulties if the prior information is poor.

The marginal decrease in information is calculated by taking the difference in information (in a thirty-two point design) between the optimal design (for the out of range parameter set) and the average design. The difference is divided by the information in the optimal design to compute the proportion that the difference represents. The implication of such a dramatic difference is that a sequential design may be doomed from the start if a design is not robust enough for the prior information.

For instance suppose the prior estimate is terrible with respect to the true parameter. This estimate is used to specify a design and an experiment is performed. Since the estimate is terrible, the design will surely be sub-optimal. Suppose first the sub-optimal design is adequate. The estimate will be updated, and a new design specified. On the other hand, suppose the information in the suboptimal design is inadequate so that the update is still in the "robust" range of the initial sub-optimal design. If this is the case, then the design will not change. However, the design won't change if it is optimal in the first place either.

In the mathematical abstraction of the system, this should not occur in high dispersion cases, since mass will move instantaneously (albeit very small) throughout the system. This is explained because the model is a parabolic partial differential equation. In a low dispersion case, when the equation becomes hyperbolic the mass will have finite velocities, and there were locations in the model where mass never appears. These locations will change as the parameters are changed. If the prior estimate causes the algorithm to choose a design which is a zero information design for the true parameters, then the update will be worthless, but there is no way beforehand to know this, and although a zero information design is useful in some sense (it tells where not to sample) it makes the next design difficult to specify. Clearly an important avenue for more study is how to ensure the first design will give some information even with terrible prior information.

In the physical system this behavior has been observed. Often in the hydraulics of wells, a radius of influence is used beyond which a particular well has no apparent effect on the system (at a particular pumping rate). If parameters concerning that wells are to be estimated, clearly a design must take samples somewhere within the radius of influence, even though the mathematical abstraction may indicate information available beyond the physical radius of influence.

An approach to robust design in the spirit of Yeh and Sun (1984) may alleviate some problems. The selection criterion is no longer simple but conceptually the approach makes more sense. It was stated earlier that reliable estimates of parameters are obtained at locations with high sensitivities. Using this a natural selection criterion was proposed and used, however the designs were not always practical. Adjusting the objective weights led to intuitively more reasonable designs, but no systematic way of selecting the weights is apparent. Rethinking the design process and considering the goal of the experiment leads to a different approach.

It was assumed implicitly that the parameters are to be obtained in order that a parametric prediction model can be run that in some sense minimizes prediction error. Using this a number of selection criteria can be proposed. For parameter identification these criteria prefer points with high sensitivity since a fairly noisy location will give a tight estimate. On the other hand, we would like our predictions to be good regardless of the reliability of our parameters. The criterion of this study gives unrealistic locations with regard to practical application unless the weights are changed.

To avoid this it is noted that prediction error is unimportant at points of disinterest, but one would like good enough estimates so that if the points become interesting later on the prediction error is still sufficiently small. A fundamental approach can be developed as follows.

Let, $c(\theta)$ be the modelled state
$\overline{c}$ be the true state

I be an identity matrix whose dimension is all computation points

L be a design matrix, same dimension as I, whose diagonal terms are indicators of whether a point, is used for estimation or not

Note: θ parameter vector

Suppose L is given (somehow). It is desired to get the best prediction error out of the system (best is smallest). Take observations on L to calculate;

$$(c(\theta) - \overline{c})^T L (c(\theta) - \overline{c})$$

This term, which will be called local error needs to be larger than some value ε, otherwise any L, which chooses all points outside of the mass envelope, will have zero local error. However, global error must be small; i.e.,

$$\min (c(\theta) - \overline{c})^T I (c(\theta) - \overline{c})$$

where, I includes all points not just design points

Thus the identification scheme for any given L is

$$\min_{\theta} (c(\theta) - \overline{c})^T I (c(\theta) - \overline{c})$$

$$\text{s.t.} \quad (c(\theta) - \overline{c})^T L (c(\theta) - \overline{c}) > \varepsilon \tag{27}$$

The design step is of course to choose the best L. This approach is essentially δ-identifiability as proposed by Yeh and Sun (1984), however the minimization is done over all space and not just observation points. The actual solution method is not pursued here, but a suggested approach is given in terms of the criterion of this study.

The objective in (27) above is like a squared sensitivity between $\hat{\theta}$ and θ^T (where $\hat{\theta}$ is the estimate and θ^T is the "true" value) with some unit perturbation length. With this in mind the criterion used here would be: choose a design that minimizes the sum of squared sensitivities, subject to the condition that the minimum cannot be too small.

In terms of the information contours as well as intuition the designs will make more sense although now one must systematically decide how to choose ε. Clearly a Monte Carlo approach can be used here, pick an ε that ensures a design remains reasonably robust for expected parameter ranges (i.e. robust on the average).

These two ideas are not pursued but are left as promising future research topics.

11. CONCLUSIONS

A maximal information criterion has been applied to an experimental design problem of groundwater hydrology. The goal was to identify reasonable sampling configurations for obtaining data to estimate model parameters. A method for identifying the parameter ranges for which a design remains optimal was presented. Several examples were presented which illustrate important implications for sampling design.

For sensitivity based designs it was found that the injection site is often a high information location. The physical explanation is that mass is localized near the source during injection and small parameter changes will make large changes in the local mass movement. The mathematical explanation is that the models are homogeneous at least to the discritization level of the solution scheme, so any high information point can be used equally well to estimate parameters. The practical implication is than any actual design should exclude the injection site to gain a more reasonable estimate that represents an average over a distance form the source.

Another implication for sampling design is that parameters when considered independently have maximum information at different locations and joint estimation requires a multi-objective approach. it was found that in two dimensional problems sensitivity to retardation is dominant in equally weighted designs.

Lastly, designs have a limited parameter range for which they are optimum. The marginal loss of information for a non-optimal design is dramatic which indicates a sequential design approach is required.

Cost variations over space were not considered, although the solution algorithm was posed in a manner that could model this. The fixed cost of a sampling network was discussed, but its solution is left for future research.

In one dimensional examples it was found that the preference of additional locations versus additional samples at existing locations depends on dispersion.

Joint configuration and scheduling was discussed but a solution was not obtained. It is believed the joint problem can be solved using a binary state dynamic programming approach.

Future research should address the robust design problem, since a robust one shot design should be preferable to a sequential approach in a geophysical problem. Construction of head sensitive and concentration sensitive designs should be considered to try to decouple the concentration sensitivity to hydraulic parameters.

12. BIBLIOGRAPHY

Ash, A. and A. Hedayat, An Introduction to Design Optimality with an Overview of the Literature, *Commun. Statis.-Theor. Meth.*, A7(14), 1295-1325, 1978.

Bear, J., *Dynamics of Fluids in Porous Media*, American Elsevier, New York, 1972.

Bear, J., *Hydraulics of Groundwater*, McGraw Hill, New York, 1979.

Carrera, J., Usnoff, E., Szidarovsky, F., A Method for Optimal Observation Network Design for Groundwater Management, *J. Hydrology*, 73, 147-163, 1984.

Dagan, G., Stochastic Modelling of Groundwater Flow by Unconditional and Conditional Probabilities. 2. The Solute Transport. *Water Resour. Res.*, 18(15), 1571-1585, 1982.

Federer, W.T. and L.N. Balaam, *Bibliography on Experiment and Treatment Design*, Pre. 1968, Edinburg: Oliver and Boyd, 1973.

Fedorov, V.V., *Theory of Optimal Experiments*, Academic Press, New York, 1972.

Hsu, N.S. and Yeh, W.W-G., Optimum Experimental Design for Parameter Identification in Groundwater Hydrology, *Water Resour. Res.*, 25(5), 1025-1040, 1989.

Kaunas, J.R. and Haimes, Y.Y., Risk Management of Groundwater Contamination in a Multiobjective Framework, *Water Resour. Res.* 22(11), 1721-1730, 1985.

Kiefer, J. and Wolfowitz, J., Equivalence of Two Extremum Problems, *Canadian J. Mathematics*, 12(3), 234-241, 1960.

Knoppman, D.S. and Voss, C.I., Behavior of Sensitivities in the One-Dimensional Advection-Dispersion Equation: Implications for Parameter Estimation and Optimal Design, *Water Resour. Res.*, 23(2), 253-272, 1987.

Knoppman, D.S. and Voss, C.I., Discrimination Among One-Dimensional Models of Solute Transport in Porous Media: Implications for Sampling Design, *Water Resour. Res.*, 24(11), 1859-1876, 1988.

Kurbrusly, C.S., Distributed Parameter System Identification: A Survey, *Int. J. Control*, 26(4), 509-535, 1977.

Kuszta, B. and Sinha, N.K., Design of Optimal Input Signals for the Identification of Distributed Parameter Systems, *Int. Journal Systems Science*, 9(1), 1-7, 1978.

Qureshi, Z.H., T.S. Ng, and G.C. Goodwin, Optimum Experimental Design for Identification of Distributed Parameter Systems, *Int. J. Control,* 31(1), 21-29, 1980.

Marsily, G. de., *Quantitative Hydrogeology,* Academic Press, New York, 1986.

McCarthy, J.M., *Optimal Pump Test Design for Parameter Estimation and Prediction in Groundwater Hydrology,* PhD Dissertation, University of California, Los Angeles, 1988.

Mehra, R.K., Optimal Input Signals for Parameter Estimation in Dynamic Systems - Survey and New Results, *I.E.E.E. Trans. Autom. Control,* 19(6), 753-768, 1974b.

Mendenhall, W., Shaffer, R. and Wackerly, D., *Mathematical Statistics with Applications,* Duxbury Press, Boston, 1986.

Meyer, P.D. and Brill, D., A Method for Locating Wells in a Groundwater Monitoring Network Under Conditions of Uncertainty, *Water Resour. Res.,* 24(8), 1277-1282, 1988.

Moody, D.W. and Maddock, T., III., A Planning Model for Preliminary Network Design, *Proceedings of International Symposium on Uncertainties in Hydrologic and Water Resource Systems, Vol. III,* Editors: Kisiel, C.C. and Duckstein, L., University of Arizona, 1972.

Naff, R.L., Yeh, J.T-C., and Kemblowski, M.W., A Note on the Recent Natural Gradient Tracer Test at the Borden Site, *Water Resour. Res.,* 24(12), 2099-2103, 1988.

Neuman, S.P., A Statistical Approach to the Inverse Problem of Aquifer Hydrology: 3. Improved Solution Method and Added Perspective, *Water Resour. Res.,* 16(2), 331-346, 1980.

Nishikawa, T., *Optimal Pump Test Design for the Parameter Identification of Groundwater Systems,* PhD Dissertation, University of California, Los Angeles, 1988.

Papadimitriou, C. and Steiglitz, K., *Combinatorial Optimization,* Prentice Hall, New Jersey, 1982.

Rafajilowicz, E., Design of Experiments for Eigenvalue Identification in Distributed-Parameter Systems, *Int. J. Control,* 34(6), 1079-1094, 1981b.

Rafajlowicz, E., Optimal Input Signals for Parameter Estimation in Linear Distributed-Parameter Systems, *Int. J. System Science,* 13(7): 799-808, 1982.

St. John, R.C. and Draper, N.R., D-Optimality for Regression Designs: A Review, *Technometrics,* 17(1), 15-23, 1979.

Sassone, P. and Schaffer, W., *Cost Benefit Analysis: A Handbook*, Academic Press, New York, 1978.

Steinberg, D.M. and Hunter, W.G., Experimental Design: Review and Comment, *Technometrics*, 26(2), 71-97, 1984.

Strecker, E.W., Chu, W-S. and Lettermaier, D.P., *Evaluation of Data Requirements for Groundwater Contaminant Transport Modelling*, Water Resources Series, Technical Report #94, University of Washington, 1985.

Tchobanoglous, G. and Schroeder, E., *Water Quality*, Addison-Wesley, 1985.

Wagner, B.J. and Gorelick, S.M., A Statistical Methodology for Estimating Transport Parameters: Theory and Applications to One-Dimensional Advective Dispersive Systems: *Water Resour. Res.*, 23(7), 1162-1174, 1987.

Wang, C., *A Three-Dimensional Finite Element Model Coupled with Parameter Identification for Aquifer Solute Transport*, PhD Dissertation, University of California, Los Angeles, 1986.

Wood, E.F., et al., *Groundwater Contamination from Hazardous Wastes*, Prentice-Hall, New Jersey, 1984.

Yeh, W.W-G. and Sun, N.Z., An Extended Identifiability in Aquifer Parameter Identification and Optimal Pumping Test Design, *Water Resour. Res.*, 20(12), 1837-1847, 1984.

Yeh, W.W-G., Review of Parameter Identification Procedures in Groundwater Hydrology: The Inverse Problem, *Water Resour. Res.*, 22(2), 95-108, 1986.

Walter, E. and Pronzato, L., Robust Experiment Design: Between Qualitative and Quantitative Identifiabilities, *Identifiability of Parametric Models*, ed. Walter, E., Pergamon Press, New York, 1987.

Sacks, J. and Ylvisaker, D., Designs for Regression Problems with Correlated Errors; Many Parameters, *The Annals of Mathematical Statistics*, 39(1), 49-69, 1968.